Psychological Medicine

cambridge.org/psm

Original Article

Cite this article: Kanayama G, Pope HG Jr, Hudson JI (2018). Associations of anabolic-androgenic steroid use with other behavioral disorders: an analysis using directed acyclic graphs. *Psychological Medicine* **48**, 2601–2608. https://doi.org/10.1017/S0033291718000508

Received: 5 December 2017
Revised: 30 January 2018
Accepted: 6 February 2018
First published online: 1 March 2018

Key words:
Anabolic-androgenic steroids; body image disorder; men; sociopathy; substance-use disorder

Author for correspondence:
Harrison G. Pope, E-mail: hpope@mclean.harvard.edu

CAMBRIDGE UNIVERSITY PRESS

Associations of anabolic-androgenic steroid use with other behavioral disorders: an analysis using directed acyclic graphs

Gen Kanayama, Harrison G. Pope Jr and James I. Hudson

From the Biological Psychiatry Laboratory, McLean Hospital/Harvard Medical School, 115 Mill St, Belmont, Massachusetts 02478, USA

Abstract

Background. Anabolic-androgenic steroid (AAS) use is known to be associated with other psychiatric disorders, such as body image disorders, conduct disorder/sociopathy, and other substance use disorders (SUD) – but the causal pathways among these conditions remain poorly delineated.

Methods. We created a directed acyclic graph to diagram hypothesized relationships among AAS use and dependence, body image disorder (BID), conduct disorder/sociopathy, and other SUD. Using proportional hazards models, we then assessed potentially causal relationships among these variables, using a dataset of 233 male weightlifters, of whom 102 had used AAS.

Results. BID and conduct disorder/sociopathy both strongly contributed to the development of AAS use, but did not appear to contribute further to the progression from AAS use to AAS dependence. Other SUD beginning prior to first AAS use – whether broadly defined or restricted only to opioids – failed to show an effect on AAS use or progression to AAS dependence. Conversely, AAS use contributed significantly to the subsequent first-time development of opioid use disorders but did not significantly increase the risk for first-time development of non-opioid SUD, taken as a whole.

Conclusions. Our analysis suggests that AAS use and other SUD are mutually attributable to underlying conduct disorder/sociopathy. SUD do not appear to represent a 'gateway' to subsequent AAS use. AAS use may represent a gateway to subsequent opioid use disorder, but probably not to other SUD.

The use of anabolic-androgenic steroids (AAS) for athletic performance or personal appearance, once restricted to elite athletes and competition bodybuilders, has evolved into a major worldwide substance use disorder over the last 30 years (Pope *et al.* 2014*b*). This development represents an emerging public-health problem since long-term AAS use can cause cardiovascular toxicity, nephrotoxicity, occasional hepatotoxicity, prolonged suppression of the hypothalamic-pituitary-testicular axis, and a range of neuropsychiatric effects (Pope *et al.* 2014*b*). Moreover, AAS can create a dependence syndrome, affecting some 30% of users (Pope *et al.* 2014*a*), in which individuals will often continue to use these drugs for years despite adverse social, psychiatric, or medical effects.

What factors increase the risk for AAS use and AAS dependence? One major factor is *male gender*. AAS use is very uncommon in women, since women rarely aspire to be musclebound, and are vulnerable to AAS-induced masculinizing effects, such as beard growth and masculinized secondary sexual characteristics. Indeed, looking at five recent studies that collectively recruited 1177 AAS users without regard to gender, the percentage of females was only 1.8% (95% confidence interval (CI): 0.8–2.7%) (Pope *et al.* 2014*a*). AAS dependence is even more uncommon in women, with only 2 AAS-dependent women among 363 cases of AAS dependence collectively reported in 10 studies worldwide (Pope *et al.* 2014*a*). For these reasons, the analysis below is restricted to males.

Accumulating evidence has suggested two major risk factors for male AAS use. The first is *body-image disorder*. Numerous studies have shown that symptoms of 'muscle dysmorphia' – a form of body dysmorphic disorder in which individuals are concerned that they are not sufficiently muscular – are common among AAS users (Pope *et al.* 1997; Olivardia *et al.* 2000; Choi *et al.* 2002; Chung, 2003; Hildebrandt *et al.* 2004; Cafri *et al.* 2005; Kanayama *et al.* 2006; Cafri *et al.* 2008; Rohman, 2009; Pope *et al.* 2012). Symptoms of muscle dysmorphia often precede the development of AAS use, but paradoxically these symptoms may also arise, or be exacerbated, after the onset of AAS use, even though the individual has often gained substantial muscle size from AAS. Thus, the causal pathways between symptoms of muscle dysmorphia and AAS use may be bidirectional.

A second major risk factor for AAS use is *conduct disorder/sociopathy*, as suggested by studies showing an association between AAS use and delinquency, violent crime, or imprisonment (Thiblin & Parlklo, 2002; Klotz *et al.* 2006; Klotz *et al.* 2007; Beaver *et al.* 2008; Klotz *et al.*

2010; Lundholm *et al.* 2010; Petersson *et al.* 2010; Skarberg *et al.* 2010; Lundholm *et al.* 2015). Again, however, the directions of causality underlying these associations are unclear. Further complicating this picture, AAS users often use other drugs (Skarberg *et al.* 2009; Dodge & Hoagland, 2011; Pope *et al.* 2012; Hallgren *et al.* 2015), and use of these other drugs may contribute to these associations. Indeed, some studies have shown that after controlling for other substance use, the association between AAS and criminality was markedly attenuated (Klotz *et al.* 2006; Lundholm *et al.* 2015). Conversely, a recent study of Swedish high school boys found that AAS users scored much higher on indices of antisocial behavior than boys using other substances but not AAS – suggesting that the association between AAS use and conduct disorder/sociopathy could not be entirely explained by use of other drugs (Hallgren *et al.* 2015). Thus, the causal pathways linking conduct disorder/sociopathy, AAS use, and other substance use are not clearly delineated. To dissect the roles of these various factors, one must consider (1) the causal pathways leading to initial AAS *use*; (2) pathways leading onward from initial AAS *use* to AAS *dependence*; and (3) pathways leading from AAS use to other disorders, such as other SUD.

In a recent study (Pope *et al.* 2012), we explored the first of these issues – causal risk factors for initial AAS use – using a 'cross-sectional cohort' design, wherein a study cohort, not selected for the outcomes of interest, is evaluated cross-sectionally, and exposures and outcomes are assessed retrospectively (Hudson *et al.* 2005). This study suggested that adolescent symptoms of muscle dysmorphia (hereafter abbreviated as simply '*body-image disorder*') and adolescent *conduct disorder* were independent causal risk factors for subsequent AAS use. However, these results left several important questions unanswered: (1) do body image disorder (BID) and conduct disorder/sociopathy not only cause AAS use, but additionally cause some men to progress from initial AAS use to AAS dependence; (2) do other SUD play a causal role in the development of AAS use and/or progression to AAS dependence (i.e. does other substance use represent a 'gateway' to AAS use); and (3) does AAS use play a causal role in the development of other SUD (i.e. does AAS use represent a 'gateway' to other substance use)? Further subdividing the latter two questions, it is notable that some studies have specifically suggested an association between opioid use and AAS use in humans (Tennant *et al.* 1988; McBride *et al.* 1996; Wines *et al.* 1999; Arvary & Pope, 2000; Kanayama *et al.* 2003) and between AAS and opioidergic mechanisms in animal models (Celerier *et al.* 2003; Peters & Wood, 2005; Nyberg & Hallberg, 2012). Thus, it is of interest to ask (a) whether opioids, in particular, play a causal role in the development of AAS use and/or dependence or (b) whether AAS use plays a causal role in the development of opioid use disorders.

Method

Source of the data

For the present analysis, we utilized data from our prior study (Pope *et al.* 2012), which assessed 233 male weightlifters, aged 18–40, of whom 102 (44%) had used AAS. These participants were recruited by advertisements in gymnasiums seeking 'men age 18 to 40 who could bench-press 275 pounds for at least one repetition,' where the bench-press requirement simply represented a device to secure reasonably experienced and accomplished weightlifters. The exposure and outcome variables of interest (e.g. AAS use) were not disclosed in the advertisements.

This study used a cross-sectional cohort design, a method whose formal properties we have previously presented (Hudson *et al.* 2005). This design identifies a dynamic cohort of individuals, drawn from a given source population, who in principle could have been enumerated in the past and followed to the present (called the 'conceptual cohort,' here representing all male weightlifters in commercial gyms capable of a 275-pound bench press during the time period beginning when the first current study participant first met these criteria). Instead of sampling from the conceptual cohort, the investigators sample in the present from individuals currently available (the 'study cohort,' here representing male weightlifters capable of a 275-pound bench press during the time of the current study and at the locations from which we recruited). With this design, estimates of effects derived from the study cohort are valid with respect to the conceptual cohort, subject to similar conditions for validity as other retrospective designs (e.g. retrospective cohort and case-control studies) (Hudson *et al.* 2005).

For all of the 233 men in the study sample, we assessed history and time of onset of lifetime *psychiatric and substance-use disorders*, using the Structured Clinical Interview for DSM-IV (First *et al.* 2001); adolescent *body-image* disorder, assessed retrospectively using a 'muscle dysmorphia' version of the Body Dysmorphic Disorder Modification of the Yale-Brown Obsessive-Compulsive Scale (Phillips *et al.* 1997; Pope *et al.* 2010); and adolescent *conduct disorder*, assessed retrospectively using the Structured Clinical Interview for DSM-IV-Child Edition (Hien *et al.* 2004). We also assessed lifetime AAS use, if any, including onset of first use and onset of AAS dependence, if it had developed – with the latter diagnosis based on criteria previously published (Kanayama *et al.* 2009*a*), and elicited using an interview module with established psychometric properties (Pope *et al.* 2010).

Directed acyclic graph analysis

For the present study, we diagrammed our hypotheses using a directed acyclic graph (DAG), which is a form of causal modeling. Technically, DAGs are a graphic representation of a recursive nonparametric structural equation model (Pearl, 1995). Causal DAGs provide a rigorous, yet intuitive, framework for examining causality (Spirtes *et al.* 1993; Greenland *et al.* 1999; Robins *et al.* 2001; Hernán & Robins, 2018), and have proven useful in many epidemiological applications (Greenland *et al.* 1999, Hernán & Robins, 2018), including psychiatric applications (Hudson *et al.* 2008; Milner *et al.* 2014). DAGs are closely related to the more familiar path diagrams (Cox & Wermuth, 1993; Pearl, 2000), but differ in that path diagrams also imply a parametric linear statistical relationship between variables. Thus, causal DAGs and the non-parametric structural equation models that they imply can be viewed as generalizations of path diagrams and corresponding linear structural equation models. A DAG implies statistical independencies and dependencies that can be read from the DAG using graphical analysis. Therefore, any statistical model for associations that fits the observed data and is consistent with the dependencies and independencies implied by the DAG can be used to estimate a causal effect of interest.

Figure 1 provides a complete DAG, diagramming all plausible causal relationships among the development of AAS use

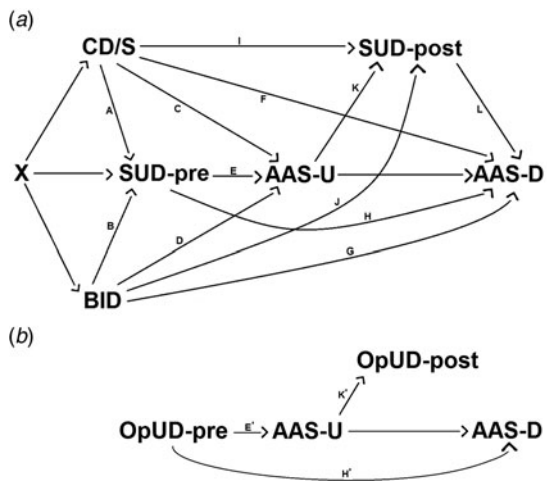

Fig. 1. (*a*). DAG 1. A complete directed acyclic graph (DAG) showing all plausible causal paths among the development of AAS use (AAS-U), the development of AAS dependence (AAS-D), adolescent body-image disorder (BID), adolescent conduct disorder/sociopathy (CD/S), other substance use disorders beginning before first AAS use (SUD-pre), and other substance use disorders beginning after first AAS use (SUD-post). Note that CD/S, BID, and SUD may, in turn, be associated via unmeasured underlying common causal factors (denoted by 'X'). See text and Table 1 for definitions of the arrows labeled 'A' to 'L'. (*b*). DAG 2. A DAG reduced from DAG 1 in panel A showing relevant paths pertaining to the relationship of opioid use disorders with onset pre-AAS (OpUD-pre) and post-AAS (OpUD-post) to AAS-U and AAS-D. See text and Table 1 for definitions of the arrows labeled 'E′′', 'H′′', and 'K′′'.

(AAS-U), the development of AAS dependence (AAS-D), adolescent BID, adolescent conduct disorder/sociopathy (CD/S), other SUD beginning before first AAS use (SUD-pre), and other SUD beginning after first AAS use (SUD-post). In the DAG, we acknowledge that BID, CD/S, and SUD may, in turn, be associated via underlying unmeasured causal factors (denoted by 'X'). Indeed, such factors are highly likely. Scientific evidence suggests that conduct disorder/sociopathy and SUD belong to a group of disorders with an underlying predisposition to 'externalizing' psychopathology (Krueger *et al.* 1998; Krueger, 1999; Kessler *et al.* 2011; Kendler & Myers, 2014). By contrast, muscle dysmorphia is likely associated with 'internalizing' psychopathology, because substantial evidence places muscle dysmorphia and other forms of body dysmorphic disorder in the 'obsessive compulsive spectrum' (Phillips *et al.* 1995; Phillips *et al.* 2010*a*, *b*), and obsessive compulsive disorder, in turn, is well established to be an internalizing disorder (Krueger *et al.* 1998; Kessler *et al.* 2011). Finally, externalizing and internalizing pathology show a substantial correlation with each other (Krueger *et al.* 1998; Krueger, 1999; Kendler & Myers, 2014).

The complete DAG contains 12 causal paths, labeled 'A' to 'L' in the figure, thus generating 12 corresponding hypotheses for testing in the present study (see Table 1). The hypotheses of primary interest here are those involving the causes and effects of AAS use and dependence (hypotheses C through G and K through L). Hypotheses C (conduct disorder/sociopathy causes AAS use) and D (BID causes AAS use) have already been supported by our prior study, as noted above (Pope *et al.* 2012). However, the other AAS-associated hypotheses remain to be tested. Specifically, what factors cause progression from initial AAS use to AAS dependence? Do other SUD lead to subsequent AAS use? Does AAS use lead to subsequent use of other drugs?

In addition, given the possible relationships between AAS and opioid use, cited above, we also tested hypotheses E, H, and K while substituting opioid use disorders for SUD overall. These hypotheses, listed as E′, H′, and K′, are diagrammed at the bottom of Fig. 1.

Statistical methods

Using our dataset, we fitted individual proportional hazards models, adjusted for self-defined race/ethnicity, study site, and birth cohort (all as previously defined (Pope *et al.* 2012)) and estimated hazard ratios corresponding 14 of the 15 paths hypothesized above. Path L could not be tested because the data were too sparse to be able to fit a model (only three men first developed other SUD after first AAS use but prior to developing AAS dependence). Note that we did not use techniques aimed at learning the structure of the causal Bayesian network represented by the DAGs (e.g. methods such as those considered in Koller and Friedman (Koller & Friedman, 2009)), nor did we attempt to fit overarching models of the entire DAG to evaluate the relative fit of competing models (e.g. by use of structural equation modeling).

We modeled BID, conduct disorder/sociopathy, and other SUD disorder as binary variables. We diagnosed other SUD using DSM-IV criteria (American Psychiatric Association, 2000); the onset of other SUD was defined as age at first occurrence of abuse of or dependence upon any non-AAS substance including alcohol, but not including nicotine. Note that predictor variables (exposures) were modeled as fixed for each subject, except that in the models for the development of AAS use, other substance use disorder was modeled as a time-varying exposure, and in the models for the development of other substance use disorder occurring after AAS use, AAS use was modeled as a time-varying exposure. Technically, the outcome variable for each analysis was the time to onset of that variable.

Given the well-established causal effect of conduct disorder/sociopathy on SUD (Weinberg & Glantz, 1999; Elkins *et al.* 2007), we performed all analyses involving time to onset of other SUD conditional on conduct disorder/sociopathy by including a term for conduct disorder/sociopathy in the proportional hazards model. For other analyses, we conditioned on other variables only to the extent that such conditioning was supported empirically by the results of the analyses (e.g. for the association of time to onset to various outcomes with body image disturbance, we conditioned on conduct disorder/sociopathy because conduct disorder/sociopathy was associated with these outcomes independently of body image disturbance, whereas for the association of these outcomes with conduct disorder/sociopathy, conditioning on body image disturbance was often not supported due to a lack of an independent effect of this variable and lack of any appreciable change in estimate when this variable was included in the model). The goal of the analysis was, for each hypothesis, to provide the best-unbiased estimate of the causal effect of a given exposure variable on the time to onset of a given outcome variable.

The assumptions of the model were those standard to proportional hazards survival models, including that censoring is non-informative and importantly that the hazard is proportional between groups over time. We checked for potential violations of the of proportional hazards assumption by inspection of log-log plots, comparison of proportional hazards model-predicted curves with Kaplan-Meier observed survival curves, and tests of weighted residuals (Grambsch & Thernau, 1994). Finally, if a given analysis failed to show a statistically significant association,

https://doi.org/10.1017/S0033291718000508 Published online by Cambridge University Press

we calculated the maximum hazard ratio that we could exclude, using a test for non-equivalence, based on the 90% CI of our measured effect sizes (Schuirmann, 1987; Berger & Hsu, 1996). This test estimates a value of the hazard ratio for which there is less than 5% probability that the true hazard ratio exceeds this magnitude; that is, a 5% error rate. All analyses were performed using Stata 14.2 (Stata Corporation, College Station, Texas), with α set at 0.05, 2-tailed.

For the association between conduct disorder/sociopathy and BID due to common unmeasured factors, we calculated the correlation coefficient (adjusted for age, race/ethnicity, and study site, using inverse probability weighting (Robbins *et al.* 2000; Hernán & Robins, 2018)).

To address the issue of multiple comparisons, we controlled for the false discovery rate (Benjamini & Hochberg, 1995) for the hypotheses tested (Table 1).

Results

The three study groups (AAS non-users, non-dependent AAS users, and dependent AAS users) showed similar demographic features but differed markedly on the candidate risk factors under analysis (Table 2). Notably, both groups of AAS users showed a high prevalence of other SUD, with some of these disorders beginning before and others after the onset of AAS use.

Table 1. Hypotheses regarding causal relationships among conduct disorder/sociopathy, body image disorder, substance use disorder, and anabolic-androgenic steroid (AAS) use and dependence

	Directed Acyclic Graph		Hazard Ratio			
Hypothesis	Letter	Path	Estimate[a]	95% CI	Excluded maximum[b]	*p* value
Conduct disorder/sociopathy causes substance use disorder pre-AAS use	A	CD/S —> SUD-pre	4.2	2.8–6.4		<0.001[c]
Body image disturbance causes substance use disorder pre-AAS use	B	BID —> SUD-pre	1.1	0.8–1.7	1.6	0.58
Conduct disorder/sociopathy causes AAS use	C	CD/S —> AAS-U	2.2	1.5–3.4		<0.001[c]
Body image disturbance causes AAS use	D	BID —> AAS-U	2.5	1.6–3.9		<0.001[c]
Pre-AAS substance use disorder[d] causes AAS use	E	SUD-pre —> AAS-U	0.8	0.5–1.4	1.3	0.47
Conduct disorder/sociopathy causes progression from AAS use to AAS dependence	F	CD/S —> AAS-D	2.0	1.0–3.9	3.5	0.04
Body image disturbance causes progression from AAS use to AAS dependence	G	BID —> AAS-D	0.9	0.4–1.8	1.6	0.72
Pre-AAS substance use disorder[d] causes progression from AAS use to AAS dependence	H	SUD-pre —> AAS-D	0.8	0.4–1.7	1.5	0.59
Conduct disorder/sociopathy causes substance use disorder post-AAS use[e]	I	CD/S —> SUD-post	4.2	2.9–6.2		<0.001[c]
Body image disturbance causes substance use disorder post-AAS use[e]	J	BID —> SUD-post	1.0	0.7–1.4	1.4	0.99
AAS use causes substance use disorder post-AAS use[e]	K	AAS-U —> SUD-post	1.6	0.9–2.8	2.5	0.14
Post-AAS substance use disorder causes progression from AAS use to AAS dependence	L	SUD-post —> AAS-D	—[f]			
Pre-AAS opioid use disorder[g] causes AAS use	E′	OpUD-pre —> AAS-U	1.3[h]	0.6–3.2	2.5	0.50
Pre-AAS opioid use disorder[g] causes progression from AAS use to AAS dependence	H′	OpUD-pre —> AAS-D	0.7	0.2–3.3	2.6	0.69
AAS use causes opioid use disorder post-AAS use[i]	K′	AAS-U —> OpUD-post	4.5	2.1–9.9		<0.001[c]
AAS use causes non-opioid substance use disorder post-AAS use[j]	K″	AAS-U —> Non-OpUD-post	1.4	0.8–2.6	2.4	0.25

AAS, anabolic-androgenic steroid.
[a]Estimate from proportional hazards model conditioned on conduct disorder/sociopathy (except when exposure under study) and further adjusted for age, race/ethnicity, study site, and birth cohort.
[b]Upper limit of hazard ratio at 5% error rate, presented for hypotheses for which there was a failure to reject the null hypothesis ($p \geqslant 0.05$).
[c]Statistically significant controlling for false discovery rate.
[d]Substance use disorder arising before first AAS use.
[e]Substance use disorder arising after first AAS use.
[f]Insufficient data to fit model.
[g]Opioid use disorder arising before first AAS use.
[h]Conditioned on body image disorder in addition to variables listed in footnote b.
[i]Opioid use disorder arising after first AAS use.
[j]Non-opioid substance use disorder arising after first AAS use.

https://doi.org/10.1017/S0033291718000508 Published online by Cambridge University Press

For the 15 hypotheses tested and shown in Table 1 (which represented the 14 *a priori* hypotheses, listed above, plus the *a posteriori* hypothesis, path K″, described below), the threshold for a 5% false discovery rate was calculated to be a p value of 0.015, based on the finding that there were five hypotheses for which the p value was less than this value. Indeed, for the tests of these five hypotheses, the p values were all <0.001 and hence would have survived even a much more conservative Bonferroni adjustment (specifically, a p value threshold of 0.0033).

As shown in Table 1, we found that conduct disorder/sociopathy showed a marked effect on the development of AAS use (path C), as had already been noted in our earlier paper (Pope *et al.* 2012). However, conduct disorder/sociopathy did not show a statistically significant effect on the progression from initial AAS use to AAS dependence (path F), although such an effect cannot be excluded with confidence (specifically, we could not exclude a hazard ratio of up to 3.5 at the 5% level of error). Similarly, adolescent BID greatly increased the risk for development of AAS use (path D) but showed little effect on the progression from AAS use to AAS dependence (path G). Specifically, as shown in Table 1, we could exclude a hazard ratio of greater than 1.9 on this variable at the 5% level of error.

Other SUD arising prior to first AAS use showed no evidence of an effect on AAS use or on the progression from AAS use to AAS dependence (paths E and H); similar results were obtained when we restricted this analysis to opioid use disorders alone (paths E′ and H′). However, we found substantial evidence that AAS use increased the risk for developing a subsequent first-onset opioid use disorder (path K′). Although AAS use did not significantly increase the risk for other SUD overall (path K), we could only exclude a hazard ratio of greater than 2.5 in this analysis. Upon seeing this result, we assessed whether AAS use exhibited an effect on all forms of first-onset SUD *excluding* opioids (path K″, which is not explicitly diagrammed in the DAG but can be easily visualized by replacing the description of path K as diagrammed with the description of path K″). This exercise yielded a nonsignificant hazard ratio of 1.4, which is slightly lower than the hazard ratio of 1.6 obtained for any SUD (Table 1).

Throughout the above analyses, no violations of the proportional hazards assumption were detected by inspection of log-log plots, comparison of model-predicted curves with Kaplan-Meier curves, or by tests of weighted residuals, except in the case of Path E (pre-AAS SUD cause AAS). For this analysis, the estimated hazard curves crossed, and thus the assumption of

Table 2. Attributes of anabolic-androgenic steroid (AAS) non-users, nondependent AAS users, and dependent AAS users

Attribute[a]	AAS non-users, $N = 131$	Nondependent AAS users, $N = 65$	Dependent AAS users, $N = 37$
Age, years	27.8 (5.8)	28.6 (5.8)	32.3 (5.8)
Self-reported race/ethnicity			
Non-Hispanic White	89 (68)	58 (89)	30 (81)
Non-Hispanic African-American	23 (18)	3 (5)	4 (11)
Hispanic White	10 (8)	–	–
Hispanic African-American	3 (2)	2 (3)	2 (5)
Asian or Pacific Islander	6 (5)	2 (3)	1 (3)
College graduate	62 (47)	21 (32)	4 (11)
Annual income > $ 50 000	30 (23)	20 (31)	12 (32)
Fat-free mass index, kg/m2	22.8 (1.9)	23.3 (2.4)	25.7 (2.7)
Years of regular weightlifting	8.4 (5.0)	9.2 (5.4)	11.7 (6.3)
Age at onset of regular weightlifting	16.1 (3.4)	16.0 (3.3)	16.5 (4.0)
Age at onset of AAS use	–	22.6 (4.7)	23.1 (5.8)
Lifetime weeks of AAS use	–	24.6 (23.7)	261 (217)
Adolescent conduct disorder	24 (18.3)	22 (33.8)	19 (51.4)
Adolescent body image disorder:			
Low (BDD-YBOCS score ⩽2)	92 (70.2)	24 (36.9)	13 (35.1)
High (BDD-YBOCS score ⩾3)	39 (29.8)	41 (63.1)	24 (64.9)
Other illicit substance abuse or dependence in lifetime:			
All substances (including alcohol)	62 (47.3)	40 (61.5)	31 (83.8)
Opioid abuse or dependence	6 (4.6)	12 (18.5)	14 (37.8)
Other illicit substance abuse or dependence prior to onset of AAS use:			
All substances (including alcohol)		31 (47.7)	17 (45.9)
Opioid abuse or dependence		3 (4.6)	2 (5.4)

AAS, anabolic-androgenic steroids; BDD-YBOCS, body dysmorphic disorder modification of the Yale-Brown Obsessive Compulsive Scale.
[a]Results are shown as mean (S.D.) for continuous variables and n (%) for discrete variables.

https://doi.org/10.1017/S0033291718000508 Published online by Cambridge University Press

proportional hazards could not be confirmed. However, the estimated risk ratio, which does not require this assumption, was the same as the estimated hazard ratio (both 0.83), and thus our estimate of association using the hazard ratio was not unduly influenced by potential violations of the proportional hazards assumption.

For the association between conduct disorder/sociopathy and body image disturbance, the correlation coefficient (95% CI) was 0.36 (0.15, 0.56), $p < 0.001$.

The above findings are reflected in a revised DAG (Fig. 2), which deletes the rejected hypothesized paths (B, E, G, H, J, and K in the upper part; E′ and H′ in the lower part). Note that we also deleted path L because we did not have enough data to test it.

Discussion

Using data from a cross-sectional cohort study of 233 male weightlifters, we tested hypothesized causal relationships, diagrammed in a DAG, among adolescent BID, adolescent conduct disorder, other SUD, development of AAS use, and progression from AAS use to AAS dependence. We confirmed our earlier findings that both adolescent BID and adolescent conduct disorder/sociopathy represent substantial causes for the development of AAS use. Moreover, within the group of AAS users, we found that a history of adolescent conduct disorder/sociopathy possibly influenced progression from AAS use to AAS dependence. However, this finding achieved a p value of 0.042, which did not survive correction for false discovery rate. Conversely, adolescent BID was not associated with an elevated hazard for progression from AAS use to AAS dependence – suggesting that adolescent BID contributes only to initial AAS use, but does not further contribute to the trajectory onwards to AAS dependence. This surprising finding contradicts our prior impression that muscle dysmorphia was an important factor in the development of AAS dependence among individuals who have already initiated AAS use (Kanayama *et al.* 2010).

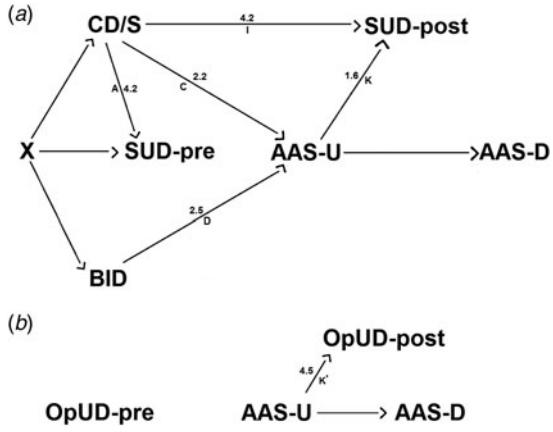

Fig. 2. (*a*). DAG 3. A revision of DAG 1 shown in Fig. 1, panel a, showing only those paths for which a significant effect was found on analysis in the present study (i.e. where the hypothesis of independence could be rejected). Although path L has been deleted in DAG 3, it should be noted that data were too sparse to fit a model to test this path and thus we could neither accept nor reject this path. (*b*). DAG 4. A revision of DAG 2 shown in Fig. 1, panel b, showing only those paths, along with the corresponding estimated hazard ratio placed above the arrows, for which a statistically significant effect was found on analysis in the present study.

Looking at participants who reported SUD other than AAS, we found that the prevalence of other SUD was lowest in AAS non-users, higher in non-dependent AAS users, and highest among dependent AAS users, with each group differing significantly from the group below. These findings are consistent with cross-sectional studies conducted elsewhere that have found a strong association between AAS use and the presence of other SUD (McCabe *et al.* 2007; Skarberg *et al.* 2009; Dodge & Hoagland, 2011; Hakansson *et al.* 2012). However, despite these strong cross-sectional associations based on lifetime prevalence, other SUD occurring before the onset of AAS use – whether defined broadly or restricted to opioid use disorders alone – exhibited no *causal* effect on the development of AAS use or progression to AAS dependence. Thus, it appears likely that AAS use and other SUD are characteristics mutually attributable to underlying conduct disorder/sociopathy, but that SUD, in and of themselves, do not predispose individuals to use AAS.

On the other hand, we found that AAS use significantly influenced the subsequent first-time development of an opioid use disorder, whereas we found little evidence that AAS use increased the risk for subsequent development of other forms of substance use collectively. The finding of an effect of AAS use on the development of opioid use disorder is congruent with an early report from our center suggesting that AAS use might represent a gateway to later opioid use disorders (Arvary & Pope, 2000). However, subsequent reports from our center (Kanayama *et al.* 2009*b*) and others (Garevik & Rane, 2010) have questioned whether this 'gateway hypothesis' is valid. The current study reopens this question, and suggests that some aspect of AAS use may increase the risk for subsequently developing opioid use disorders. One might speculate that young AAS users may get introduced to opioids and possibly other drugs after meeting AAS-using peers., Another possibility is that individuals may first learn to use a needle via AAS injections (Kanayama *et al.* 2003), and this might represent a stepping-stone to other injection drug use.

Our findings are subject to several limitations, many of which are inherent to cross-sectional studies that rely on retrospective reporting. Although we advertised for study participants without revealing the exposure variables (e.g. conduct disorder) and outcome variables (e.g. AAS use and dependence) of interest, participants may nevertheless have self-selected in some manner associated with these variables. However, any such selection bias was mitigated by our use of adjustments based on covariates, so that our findings need only be unbiased within strata of these covariates. Thus, we assumed only that the sample was representative of the underlying source population of experienced weightlifters when adjusted for covariates. Our results may also have been influenced by response bias if participants reported exposures in a differential manner (e.g. if AAS users were more likely to recall or disclose other drug use than non-users). Observer bias might also have occurred if we knew the group status of a participant in advance (i.e. AAS user or non-user). Although we attempted to maintain blindness to group status during the evaluations, some men were so obviously muscular that they were visibly AAS users. A detailed discussion of these and other possible limitations is provided in our earlier paper (Pope *et al.* 2012).

Caution should also be exercised when considering pathways that we rejected for lack of a significant association. In tests for independence, the null hypothesis is that there is an association (analogous to the null hypothesis of non-equivalence in a clinical trial of a novel treatment *v.* a standard treatment). Thus, one can never fully reject the null hypothesis, but rather only reject the

https://doi.org/10.1017/S0033291718000508 Published online by Cambridge University Press

hypothesis that the association is larger than a certain value, thereby concluding it is within a certain range. With our data, as seen in Table 1, there were four estimates of the hazard ratio that were between 0.8 and 1.0 and for which we could exclude hazard ratios as high as 1.3–1.6 (Paths B, E, G, H, J). Therefore, in these cases, we could reasonably exclude strong effects. However, we could not exclude hazard ratios of greater than 2.4 in five cases (Paths F, K, E′, H′, and K″), and thus we may have failed to detect important effects.

Finally, the validity of this analysis depends critically on the underlying assumptions of the proposed initial DAG. However, for this analysis, there were no obvious alternative causal structures for the variables measured in the DAG, and no other obvious factors that were omitted and that would threaten the validity of the findings. Unmeasured confounding is always a threat to validity, but the presence of conduct disorder/sociopathy and BID in the DAG, representing conditions that can serve as partial proxies for externalizing and internalizing psychopathology, respectively, mitigates the threat that the findings might be highly influenced by unmeasured underlying psychopathological factors. Another threat to validity arises from conditioning on variables that may induce additional associations, via unmeasured variables, between the predictor and the outcome. This threat, technically termed 'collider-stratification bias' (Greenland, 2003), would only occur in our analyses conditioning on conduct disorder/sociopathy if there were factors (not considered in Fig. 1*a*) that (a) caused conduct disorder/sociopathy, (b) caused the outcome of interest, and (c) were independent of the common factors for conduct disorder/sociopathy and the predictor. Not only is there no evidence for the presence of such factors, but they would also need to be strongly associated with both predictor and outcome to influence appreciably the estimates of the hazard ratios. Furthermore, the direction of any collider-stratification bias would be towards the null, thus making the estimates more conservative.

In summary, although our findings would benefit from cross-validation in a new sample, and will require exploration in further studies, they appear unlikely to have been seriously biased by the various limitations summarized above. In any case, it is clear that the associations among AAS use and dependence, conduct disorder/sociopathy, BID, and other SUD represent a complicated network of causal relationships that can only be disentangled by future longitudinal studies using rigorous methods of causal inference.

Acknowledgements. The authors would like to acknowledge Chris Perriello for technical assistance on the manuscript.

References

American Psychiatric Association (2000) *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision* (*DSM-IV-TR*). Washington: American Psychiatric Association.

Arvary D and Pope HG (2000) Anabolic-androgenic steroids as a gateway to opioid dependence. *New England Journal of Medicine* **342**, 1532.

Beaver KM, Vaughn MG, Delisi M and Wright JP (2008) Anabolic-androgenic steroid use and involvement in violent behavior in a nationally representative sample of young adult males in the United States. *American Journal of Public Health* **98**, 2185–2187.

Benjamini Y and Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289–300.

Berger RL and Hsu JC (1996) Bioequivalence trials, intersection-union tests and equivalence confidence sets. *Statistical Science* **11**, 283–302.

Cafri G, Olivardia R and Thompson JK (2008) Symptom characteristics and psychiatric comorbidity among males with muscle dysmorphia. *Comprehensive Psychiatry* **49**, 374–379.

Cafri G, Thompson JK, Ricciardelli L, McCabe M, Smolak L and Yesalis C (2005) Pursuit of the muscular ideal: physical and psychological consequences and putative risk factors. *Clinical Psychology Review* **25**, 215–239.

Celerier E, Yazdi MT, Castane A, Ghozland S, Nyberg F and Maldonado R (2003) Effects of nandrolone on acute morphine responses, tolerance and dependence in mice. *European Journal of Pharmacology* **465**, 69–81.

Choi PY, Pope HG and Olivardia R (2002) Muscle dysmorphia: a new syndrome in weightlifters. *British Journal of Sports Medicine* **36**, 375–376; discussion 377.

Chung B (2003) Muscle dysmorphia in weightlifters. *British Journal of Sports Medicine* **37**, 280–281.

Cox DR and Wermuth N (1993) Linear dependencies represented by chain graphs. *Statistical Science* **8**, 204–218.

Dodge T and Hoagland MF (2011) The use of anabolic androgenic steroids and polypharmacy: a review of the literature. *Drug and Alcohol Dependence* **114**, 100–109.

Elkins IJ, McGue M and Iacono WG (2007) Prospective effects of attention-deficit/hyperactivity disorder, conduct disorder, and sex on adolescent substance use and abuse. *Archives of General Psychiatry* **64**, 1145–1152.

First M, Spitzer R, Gibbon M and Williams J (2001) *Structured Clinical Interview for DSM-IV Axis I Disorders — Non-Patient Edition*. New York: Biometrics Research Department, New York State Psychiatric Institute.

Garevik N and Rane A (2010) Dual use of anabolic-androgenic steroids and narcotics in Sweden. *Drug and Alcohol Dependence* **109**, 144–146.

Grambsch P and Thernau T (1994) Proportional hazards test and diagnostics based on weighted residuals. *Biometrika* **81**, 515–526.

Greenland S (2003) Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology* **14**, 300–306.

Greenland S, Pearl J and Robins JM (1999) Causal diagrams for epidemiologic research. *Epidemiology* **10**, 37–48.

Hakansson A, Mickelsson K, Wallin C and Berglund M. (2012) Anabolic androgenic steroids in the general population: user characteristics and associations with substance use. *European Addiction Research* **18**, 83–90.

Hallgren M, Pope HG, Kanayama G, Hudson JI, Lundin A and Kallmen H (2015) Anti-social behaviors associated with anabolic-androgenic steroid use among male adolescents. *European Addiction Research* **21**, 32132–6.

Hernán MA and Robins JM (2018) *Causal Inference*. Boca Raton, Florida: Chapman & Hall/CRC.

Hien D, Matzner F, First M, Spitzer R, Gibbon M and Williams J (2004) *Structured Clinical Interview for DSM-IV - Child Edition, 2004*. Available at http://www.columbia.edu/~jam119/KidSkid.htm (accessed 20 February 2018).

Hildebrandt T, Langenbucher J and Schlundt DG (2004) Muscularity concerns among men: development of attitudinal and perceptual measures. *Body Image* **1**, 169–181.

Hudson JI, Javaras KN, Laird NM, VanderWeele TJ, Pope HG and Hernan MA (2008) A structural approach to the familial coaggregation of disorders. *Epidemiology* **19**, 431–439.

Hudson JI, Pope HG and Glynn RJ (2005) The cross-sectional cohort study: an underutilized design. *Epidemiology* **16**, 355–359.

Kanayama G, Barry S, Hudson JI and Pope HG (2006) Body image and attitudes toward male roles in anabolic-androgenic steroid users. *American Journal of Psychiatry* **163**, 697–703.

Kanayama G, Brower KJ, Wood RI, Hudson JI and Pope HG (2009*a*) Issues for DSM-V: clarifying the diagnostic criteria for anabolic-androgenic steroid dependence. *American Journal of Psychiatry* **166**, 642–645.

Kanayama G, Brower KJ, Wood RI, Hudson JI and Pope HG (2010) Treatment of anabolic-androgenic steroid dependence: emerging evidence and its implications. *Drug and Alcohol Dependence* **109**, 6–13.

Kanayama G, Cohane GH, Weiss RD and Pope HG (2003) Past anabolic-androgenic steroid use among men admitted for substance abuse treatment: an underrecognized problem? *Journal of Clinical Psychiatry* **64**, 156–160.

Kanayama G, Hudson JI and Pope HG (2009*b*) Features of men with anabolic-androgenic steroid dependence: a comparison with nondependent AAS users and with AAS nonusers. *Drug and Alcohol Dependence* **102**, 130–137.

Kendler KS and Myers J (2014) The boundaries of the internalizing and externalizing genetic spectra in men and women. *Psychological Medicine* **44**, 647–655.

Kessler RC, Ormel J, Petukhova M, McLaughlin KA, Green JG, Russo LJ *et al.* (2011) Development of lifetime comorbidity in the World Health Organization world mental health surveys. *Archives of General Psychiatry* **68**, 90–100.

Klotz F, Garle M, Granath F and Thiblin I (2006) Criminality among individuals testing positive for the presence of anabolic androgenic steroids. *Archives of General Psychiatry* **63**, 1274–1279.

Klotz F, Petersson A, Hoffman O and Thiblin I (2010) The significance of anabolic androgenic steroids in a Swedish prison population. *Comprehensive Psychiatry* **51**, 312–318.

Klotz F, Petersson A, Isacson D and Thiblin I (2007) Violent crime and substance abuse: a medico-legal comparison between deceased users of anabolic androgenic steroids and abusers of illicit drugs. *Forensic Science International* **173**, 57–63.

Koller D and Friedman M (2009) *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, Massachusetts: MIT press.

Krueger RF (1999) The structure of common mental disorders. *Archives of General Psychiatry* **56**, 921–926.

Krueger RF, Caspi A, Moffitt TE and Silva PA (1998) The structure and stability of common mental disorders (DSM-III-R): a longitudinal-epidemiological study. *Journal of Abnormal Psychology* **107**, 216–227.

Lundholm L, Frisell T, Lichtenstein P and Langstrom N (2015) Anabolic androgenic steroids and violent offending: confounding by polysubstance abuse among 10365 general population men. *Addiction* **110**, 100–108.

Lundholm L, Kall K, Wallin S and Thiblin I (2010) Use of anabolic androgenic steroids in substance abusers arrested for crime. *Drug and Alcohol Dependence* **111**, 222–226.

McBride AJ, Williamson K and Petersen T (1996) Three cases of nalbuphine hydrochloride dependence associated with anabolic steroid use. *British Journal of Sports Medicine* **30**, 69–70.

McCabe SE, Brower KJ, West BT, Nelson TF and Wechsler H (2007) Trends in non-medical use of anabolic steroids by U.S. College students: results from four national surveys. *Drug and Alcohol Dependence* **90**, 243–251.

Milner A, Page A and LaMontagne AD (2014) Cause and effect in studies on unemployment, mental health and suicide: a meta-analytic and conceptual review. *Psychological Medicine* **44**, 909–917.

Nyberg F and Hallberg M (2012) Interactions between opioids and anabolic androgenic steroids: implications for the development of addictive behavior. *International Review of Neurobiology* **102**, 189–206.

Olivardia R, Pope HG and Hudson JI (2000) Muscle dysmorphia in male weightlifters: a case-control study. *American Journal of Psychiatry* **157**, 1291–1296.

Pearl J (1995) Causal diagrams for empirical research. *Biometrika* **82**, 669–688.

Pearl J (2000) *Causality*. Cambridge, England: Cambridge University Press.

Peters KD and Wood RI (2005) Androgen dependence in hamsters: overdose, tolerance, and potential opioidergic mechanisms. *Neuroscience* **130**, 971–981.

Petersson A, Bengtsson J, Voltaire-Carlsson A and Thiblin I (2010) Substance abusers' motives for using anabolic androgenic steroids. *Drug and Alcohol Dependence* **111**, 170–172.

Phillips KA, Hollander E, Rasmussen SA, Aronowitz BR, DeCaria C and Goodman WK (1997) A severity rating scale for body dysmorphic disorder: development, reliability, and validity of a modified version of the Yale-Brown Obsessive Compulsive Scale. *Psychopharmacology Bulletin* **33**, 17–22.

Phillips KA, McElroy SL, Hudson JI and Pope HG (1995) Body dysmorphic disorder: an obsessive-compulsive spectrum disorder, a form of affective spectrum disorder, or both? *Journal of Clinical Psychiatry* **56**(Suppl. 4), 41–51; discussion 52.

Phillips KA, Stein DJ, Rauch SL, Hollander E, Fallon BA, Barsky A *et al.* (2010*a*) Should an obsessive-compulsive spectrum grouping of disorders be included in DSM-V? *Depression and Anxiety* **27**, 528–555.

Phillips KA, Wilhelm S, Koran LM, Didie ER, Fallon BA, Feusner J *et al.* (2010*b*) Body dysmorphic disorder: some key issues for DSM-V. *Depression and Anxiety* **27**, 573–591.

Pope HG, Gruber AJ, Choi P, Olivardia R and Phillips KA (1997) Muscle dysmorphia. An underrecognized form of body dysmorphic disorder. *Psychosomatics* **38**, 548–557.

Pope HG, Kanayama G, Athey A, Ryan E, Hudson JI and Baggish A (2014*a*) The lifetime prevalence of anabolic-androgenic steroid use and dependence in Americans: current best estimates. *American Journal on Addictions* **23**, 371–377.

Pope HG, Kanayama G and Hudson J (2012) Risk factors for illicit anabolic-androgenic steroid use in male weightlifters: a cross-sectional cohort study. *Biological Psychiatry* **71**, 254–261.

Pope HG, Kean J, Nash A, Kanayama G, Samuel DB, Bickel WK *et al.* (2010) A diagnostic interview module for anabolic-androgenic steroid dependence: preliminary evidence of reliability and validity. *Experimental and Clinical Psychopharmacology* **18**, 203–210.

Pope HG, Wood R, Rogol A, Nyberg F, Bowers L and Bhasin S (2014*b*) Adverse health consequences of performance-enhancing drugs: an endocrine society scientific statement. *Endocrine Reviews* **35**, 341–375.

Robbins J, Hernán MA and Brumback B (2000) Marginal structural models and causal inference in epidemiology. *Epidemiology* **11**, 550–560.

Robins JM, Smoller JW and Lunetta KL (2001) On the validity of the TDT test in the presence of comorbidity and ascertainment bias. *Genetic Epidemiology* **21**, 326–336.

Rohman L (2009) The relationship between anabolic androgenic steroids and muscle dysmorphia: a review. *Eating Disorders* **17**, 187–199.

Schuirmann DJ (1987) A comparison of the 2 one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics* **15**, 657–680.

Skarberg K, Nyberg F and Engstrom I (2009) Multisubstance use as a feature of addiction to anabolic-androgenic steroids. *European Addiction Research* **15**, 99–106.

Skarberg K, Nyberg F and Engstrom I (2010) Is there an association between the use of anabolic-androgenic steroids and criminality? *European Addiction Research* **16**, 213–219.

Spirtes P, Glymour C and Scheines R (1993) *Causation, Prediction and Search*. Springer-Verlag: New York.

Tennant F, Black DL and Voy RO (1988) Anabolic steroid dependence with opioid-type features. *New England Journal Of Medicine* **319**, 578.

Thiblin I and Parlklo T (2002) Anabolic androgenic steroids and violence. *Acta Psychiatrica Scandanavica Supplementum* **412**, 125–128.

Weinberg NZ and Glantz MD (1999) Child psychopathology risk factors for drug abuse: overview. *Journal of Clinical Child Psychology* **28**, 290–297.

Wines Jr. JD, Gruber AJ, Pope HG and Lukas SE (1999) Nalbuphine hydrochloride dependence in anabolic steroid users. *American Journal on Addictions* **8**, 161–164.