PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Criminality in males with Klinefelter syndrome and XYY syndrome - a cohort study
AUTHORS	Kirstine Stochholm, Anders Bojesen, Anne Skakkebæk Jensen, Svend Juul, Claus Højbjerg Gravholt

VERSION 1 - REVIEW

REVIEWER	David J Handelsman Professor of Reproductive Endocrinology & Andrology Director, ANZAC Research Institute, University of Sydney & Head, Andrology Department, Concord Hospital
REVIEW RETURNED	27/11/2011

THE STUDY	These is a major ascertainment bias due to underdiagnosis which needs to be better explored.
GENERAL COMMENTS	Historically, excess criminality among aneuploid men was reported from studies in the 1960's-70s. These studies had weak designs by modern standards using small, convenience samples that were marred by major ascertainment bias low socio-economic, institutionalised men. This created a still prevalent belief that aneuploid men had inherent genetically determined criminality. In this context this manuscript is an interesting and important study using the valuable opportunities provided by the Danish centralised and linked national health and criminal registers. The study proves beyond reasonable doubt, in these two large cohorts of aneuploid men with KS or 47 XYY with multiple, well-matched controls, there is an age-dependent and progressive excess of criminality. This excess appears to be strongly and largely, but not fully, determined by concomitant low socio-economic status. The excess criminality that precedes diagnosis, together with the similar excesses in both aneuploid groups in a well-matched cohort design, strengthens the findings. These findings raise important issues about whether early diagnosis and treatment (behavioural and possibly hormonal) might ameliorate such social determinants of criminality. Nevertheless the study has important flaws not fully acknowledged and which question the extrapolatbility of these findings to all aneuploid men. These limitations could be better integrated into the data analysis and discussion to further enhance the paper, which will be influential and widely cited in medicine and criminology. 1. Both Danish and UK linkage studies established clearly that 75% of men with 47 XXY and possibly higher proportion of men with 47 XYY are never diagnosed during life. This applies even in countries with national health schemes that eliminate poor access to health care as a reason for underdiagnosis. Hence both cohorts are only a

	small minority of the contemporaneous birth cohorts thereby introducing a major bias. As the effect size observed is modest, this ascertainment bias might have a relatively large impact. In addition to stronger caveats, some sensitivity analyses would be useful – can the authors define the reasons for cytogenetic diagnosis and stratify those for whom the diagnosis had no plausible confounding linkage with criminality? What impact would there be if the undiagnosed had only, say, 50% (or 25%, or 75%) as much excess criminality? Can the authors estimate quantitatively how much of a reduction in criminality could still be detected if the undiagnosed men were included in a simulated analysis? 2. It is not clear why the controls were matched on age and birth cohort but not on socio-economic factors when the latter proved so critical to the ultimate findings. It might have been argued to use the covariate adjustment approach so each covariate could be evaluated individually. However as all covariates were adjusted en bloc, this does not seem logical. 3. The use of “fatherhood” for socio-economic matching is a strange choice when sterility is an almost invariable feature of KS. Presumably the KS men have step-children or progeny via artificial insemination. Some explanation is required for the choice and its analytical consequences. 4. To interpret the reduction in driving offenses and drug-related crime in KS men it would be interesting to know if this excess could be explained by adjustment for having a driving licence and/or car ownership – are such data available for linkage studies? 5. There are logical deductions regarding the role of testosterone treatment that are not mentioned in Discussion. Firstly, criminality was equally or greater in the 47 XYY group (neither testosterone deficient nor treated with testosterone) in contrast to the 47 XXY men who are testosterone deficient and treated with testosterone from diagnosis after adolescence onwards. Hence neither testosterone deficiency nor treatment are likely to have contributed to the excess criminality. Secondly, the excess criminality also preceded diagnosis (and therefore treatment) so it is an unlikely explanation of the excess criminality. Thirdly, the conventional testosterone replacement doses used in treatment are not sufficient to cause the behavioural effects observed in a small minority of men supra-physiological testosterone doses. Furthermore, as compliance with testosterone treatment is also often inadequate, the net doses of testosterone are often lower than even replacement requirements.

REVIEWER	Nicole Tartaglia, MD Developmental-Behavioral Pediatrician Assistant Professor, Department of Pediatrics University of Colorado School of Medicine Child Development Unit, Children's Hospital Colorado
REVIEW RETURNED	28/12/2011

GENERAL COMMENTS	Thank you for inviting me to review the article titled “Criminality in males with Klinefelter syndrome and XYY syndrome.” While the study and statistical analysis are quite comprehensive, I do not see the value of adding this paper to the medical literature. As part of a community of researchers working to increase information about the behavioral, psychological, and medical features of these genetic conditions, the goal of papers should be to contribute new information to the literature. As the authors point out in their

	introduction, papers about increased rates of criminality and other aberrant behaviors in the sex chromosome aneuploidy conditions were first published in the 60's and 70's. These papers were met with significant criticism due to concerns about ascertainment bias and lack of accounting for other confounding factors. While these authors may use a new and more comprehensive registry and new statistical methods, the conclusions are similar to those of the papers published 50+ years ago, do not offer a new explanation for the findings (although some suggestions based on the work of others are proposed in the discussion), and I do not believe add significant information to the current literature on these conditions. While the authors try to explain and justify the paper and the reason for the study, the end result still leaves the simple message of "males with Klinefelter syndrome or XYY syndrome are more likely to be sex offenders or have other criminal behavior." This conclusion is not at all helpful to the population being studied, as it increases the stigmatization that is already present, does not provide an explanation for the findings, and will cause more parents and individuals with these conditions to be feel shameful of the diagnosis. Negative papers with these themes that cannot offer solutions also discourage clinical research participation due to concerns of further stigmatization, lead to distrust of researchers, and ultimately decrease research funding for these conditions. Furthermore, with the internet, the medical literature is no longer solely accessible to medical professionals – individuals with XXY and XYY, and their families, teachers, peers, and employers all have access to these results, which can cause further discrimination and stigmatization. While this list of consequences may seem extreme, I work closely with a large population of patients and families affected by XYY and XXY, and I have seen and heard these comments over and over, have seen research programs fail due to similar situations, and have seen the consequences of misperceptions about the condition. The field is making progress toward eliminating these biases and stigmatizing features, and publication of a paper about criminality without new solutions undermines this progress. If the goal of the research is to prove an association between a genetic condition and criminal behaviors, it is critical that the authors control for all other factors that may be contributing to the criminal behavior. As they point out (in part), the psychological factors associated with XXY and XYY include increased rates of learning and cognitive disabilities, language disorders, social cognitive deficits and social difficulties, executive function impairments, and other emotional disorders which may affect their behavior and subsequently their likelihood of being involved in and convicted of criminal acts. If the authors have a mechanism of controlling for these factors and then showing that there continues to be increased rates of criminality in XXY and XYY, this would be an important feature of the genetic condition. However, I doubt that there are significant differences in criminal convictions between those with similar cognitive levels and decision making abilities. The population of individuals convicted of crimes is not comparable to the average population (even a sample of 10 people born in the same time period), and thus I believe it does an injustice to a population of individuals with XXY and XYY (who have quite wide variability in behaviors and cognitive abilities) to associate the genetic condition to the criminal behaviors without directly controlling for these factors. The significant decreases in the HR values reported by the authors when controlling for the limited socio-economic factors they were able to control for (that are much less direct indicators of the psychological factors listed above) suggest

	that there would be little to no significance when these were controlled directly. Suggestions would be to make major revisions with a different focus and a different title – perhaps taking an approach such as "improved socioeconomic factors are associated with fewer convictions of crime in XXY and XYY." However, overall I do not support publication for all the reasons described above.

VERSION 1 – AUTHOR RESPONSE

We thank the reviewers for their constructive and spirited comments to our paper. We have dealt with all questions, comments and suggestions as can be seen below in detail. In the manuscript the changes we have made are marked with yellow.
We think that the manuscript has been improved in this process and we hope that it is now acceptable for BMJ Open. Especially, we now include a sensitivity analysis that underscores our results and emphasizes that even if we were able to study all males with Klinefelter syndrome and 47,XYY, we would still find an increased rate of criminality.
We therefore resubmit our paper: "Criminality in males with Klinefelter and XYY syndrome – a cohort study" by Kirstine Stochholm, Anders Bojesen, Anne S. Jensen, Svend Juul, Claus Højbjerg Gravholt.

Yours sincerely

Claus Højbjerg Gravholt, consultant, MD, PhD,
Department of Endocrinology and Internal Medicine
Århus Sygehus
Århus University Hospital
8000 Århus C

Response to David J Handelsman

1. We already discuss at length that non-diagnosis and delay in diagnosis is a significant problem with both KS and 47,XYY (as it also is with a wide range of other rare diseases). We can of course only state for sure that the current picture of criminality relates to the studied population, but we tentatively suggest that it might extend to other probably similar populations in other countries. However, what about the large group of undiagnosed cases? As suggested by the reviewer, we now include a sensitivity analysis. In the statistics section we have included the following: "To examine a potential bias associated with undiagnosed KS and 47,XYY cases we performed a sensitivity analysis, assuming that the risk of conviction among undiagnosed cases is smaller than the risk observed

among diagnosed cases, and we applied the statistical uncertainty from the observed data expressed by the standard error of the ln(HR) estimate". We made two sensitivity analyses – one where we assumed that the rate of criminality would be half the observed among the cases, and another analysis where we assumed that the rate of criminality would be similar to that of the controls. We have included this information in the Discussion, which have allowed us a more nuanced discussion of the results. In essence these analyses shows that even with a much lower crime rate among the undiagnosed cases, one would still observe an elevated rate of crime among a totally unbiased and diagnosed population of both KS and 47,XYY. It now reads: "We then performed a sensitivity analysis to examine a potential bias arising if the severity of the syndrome affects both the risk of conviction and the probability of being diagnosed. In one analysis we assumed that the excess hazard among undiagnosed cases was half the excess hazard seen among diagnosed cases, and we applied the statistical uncertainty from the observed data. In another similar analysis we assumed that the excess hazard among undiagnosed cases was similar to that in the background population. For KS we assumed that 25% of all cases had been diagnosed. In the sensitivity analyses the hazard ratios were reduced, but still significantly elevated for all convictions (excluding traffic offenses), and for sexual abuse, burglary, and arson (Supplementary Table 2). For 47,XYY we observed a similar pattern. Here we assumed that 15% of all cases had been diagnosed. In other words, it is highly likely that the crime rate would remain significantly increased among an entirely unbiased population of both KS and 47,XYY with complete diagnosis of all cases.".

2. As mentioned in the Discussion, we matched patients with up to 100 age and calendar matched controls from the general population. The reviewer asks why we did not match on socio-economic factors as well. There is a considerable body of research on matching, showing that it is dangerous and sometimes outright wrong to match on factors that, like in the current setting, may be causally involved in the chain of events leading to increased criminality. This can lead to overmatching 1 – "matching on factors that are affected by the study exposure or disease (i.e being KS or control) is almost never warranted and is potentially capable of biasing study results beyond any hope of repair. It is therefore crucial to understand the nature of such overmatching and why it needs to be avoided". And since we have previously shown that having KS is related to poorer socio-economic outcome, one cannot match on such factors since we do not know the precise causal relationship between these factors, the diseases (i.e. KS or 47,XYY) and criminality. Such an approach may even lead to "an irreparable form of selection bias" 1. In response to your comment we have expanded the paragraph in the Discussion now reading: "Indeed, matching on socio-economic factors would likely lead to overmatching – which "is potentially capable of biasing study results beyond any hope of repair"

3. You are of course right with regard to "fatherhood" in as much as most males with KS are infertile and do not normally become biological fathers (although we have 3 males in our clinic which presumably fathered their own children (we have not, so far, formally excluded the milkman as being the father!)). However, since many of the patients we see in our outpatient clinic do become fathers after donor insemination we therefore deemed it interesting to study this parameter. And as can be seen in our previous publication on socio-economy, surprisingly many with KS (26.2% compared with 61.9% in the background population) do actually become fathers, although we are not able to discern in the registry between biological and donor-insemination fatherhoods. And since we actually do not match on fatherhood, but only control or adjust in the subsequent analyses for this factor, we would like to keep "fatherhood" in the analyses. In our opinion, this factor proves that many KS actually do quite well in society, as "fatherhood" can be seen as a stabilizing factor, possibly even more so when one has to go through the hassles of donor insemination of semen, being an obstacle most men do not have to go through.

4. Unfortunately, we do not have data on who hold a driving license and who do not. We did actually speculate on this matter ourselves, because we also realized that it would have been interesting to know if not having a drivers license was the cause for the fewer driving offenses. However, even if there was no difference between KS and 47,XYY and the controls in having a drivers license, the poorer economical situation for the cases could lead to fewer KS and 47,XYY owning a car, which

could then explain the difference in driving offenses. Suffice to say, we do not have the data, and even if we had, we would still lack data on who own a car and who do not own a car. We have not expanded further on this issue.
5. We thank the reviewer for the constructive and stimulating comments regarding the impact of hypogonadism and testosterone treatment. In response to the comments we have expanded the discussion of testosterone supplementation and the latter part of it now reads like this: “We note that the pattern of criminality in 47,XYY, who have a normal testosterone production, was equal or higher than among KS, and furthermore that criminality among KS was elevated even before diagnosis and thus before commencement of supplementation with testosterone, making it unlikely that testosterone supplementation is causally involved in the excess criminality in KS. In addition in many KS males conventional testosterone supplementation is often not sufficient and many KS males are also not compliant, at least not all the time, resulting in hypotestosteronemia, elevated LH and diseases, symptoms and signs related to hypogonadism”.

Response to Nicole Tartaglia

We thank the reviewer for her general comments and can only conclude, as physicians caring for hundreds of patients with sex chromosome disorders through many years, being involved in research and the generation of information material for patients and relatives alike, that we apparently could not disagree more with the reviewer on the need for the publication of this paper. Before submission of our paper we had a general discussion of the results of our study within the group of authors and with other colleagues and everybody agreed that it would be unethical not to publish the data! We have a similar situation with patients with schizophrenia which also have an appallingly high rate of criminality. There is a large body of evidence to support this and this knowledge has led to the development of tools to estimate risk of criminality , as well as an increased focus on improved medicinal treatment of those with an estimated high risk of future criminality 2. In other words, the knowledge of an increased rate of criminality has been used in a constructive way to improve patient care.

1. The reviewer, we believe, misses an important point in reading our paper. No matter how one goes about the data, no matter what factor one could adjust for in subsequent analyses and no matter how benevolent one would view the results, criminality is seriously increased in several subtypes of crime among both KS and 47,XYY males! This fact is a good and constructive basis for further research and development of new treatment strategies within the realm of cognitive therapy, school support, drug treatment, etc. We have expanded on the issue of non-diagnosis in response to Dr. Handelsman and have included a sensitivity analysis and this analysis support the notion that even if there were a much lower crime rate among non-diagnosed cases, the overall picture would still be one of an increased crime rate in a totally unbiased group of both KS and 47,XYY (please see above).
2. Regarding control of factors that may or may not contribute to criminality: As mentioned in the Discussion, we matched patients with up to 100 age and calendar matched controls from the general population. The reviewer asks why we did not match on socio-economic factors as well. There is a considerable body of research on matching, showing that it is dangerous and sometimes outright wrong to match on factors that, in the current setting, may be causally involved in the chain of events leading to increased criminality. This can lead to overmatching 1 – “matching on factors that are affected by the study exposure or disease (i.e being KS or control) is almost never warranted and is potentially capable of biasing study results beyond any hope of repair. It is therefore crucial to understand the nature of such overmatching and why it needs to be avoided”. And since we have previously shown that having KS (as does males with 47,XYY) is related to poorer socio-economic outcome, one cannot and should not match on such factors since we do not know the precise causal relationship between these factors, the diseases (i.e. KS or 47,XYY) and criminality. Such an

approach may even lead to “an irreparable form of selection bias” 1. In response to your comment we have expanded the paragraph in the Discussion now reading: “Indeed, matching on socio-economic factors would likely lead to overmatching – which “is potentially capable of biasing study results beyond any hope of repair”. In addition, we would like to add the highly significant finding of increased risk of being convicted of arson and sexual abuse, which we find extremely interesting and of course also disturbing. We note that this increased risk is not decreased substantially by adjusting for socio-economic factors. We simply fail to understand that the reviewer do not find this worthy of further study! Further study of course only being possible if one acknowledges that there might be a clinical problem.
3. The suggestion to control for or match on cognitive disability, language disorders, social cognitive deficits etc is of course impossible to implement, since only a minuscule fraction of both cases and controls would have been subjected to such tests. We have not added further on this matter in the Discussion. And we again refer to the risk of ruining a dataset if one matches patients and controls on too many important factors likely to influence the course of our lives.
4. We have kept the title as is, since we believe that it accurately reflects the study results.

Reference List

1. Rothman KJ, Greenland S, Lash TL. Design strategies to improve study accuracy. In Rothman KJ, Greenland S, Lash TL: Modern Epidemiology 2008; 3rd ed:168-182.
2. Topiwala A,Fazel S. The pharmacological management of violence in schizophrenia: a structured review. Expert Rev Neurother 2011; 11:53-63.

VERSION 2 – REVIEW

REVIEWER	Professor David J Handelsman Director, ANZAC Research Institute University of Sydney Sydney NSW 2139 Australia No competing interests in this paper
REVIEW RETURNED	01/02/2012

RESULTS & CONCLUSIONS	The revision responds well to most concerns raised except for the SE matching issue. Just as the authors are fearless in studying a contentious topic, while accepting their caveats on over-matching (although recognising the choice of variables as being (over)matched is subjective), they should be equally fearless in analysis of whether SE status "explains" their striking findings. There is no penalty for peeking at this important issue and exploratory SE matching remains of strong interest without undermining the originality and impact of their work. Such further analysis is strongly encouraged.