Gander Research

ganderresearch.org

True or false, or somewhere between?

A review of the high-quality studies on the prevalence of false sexual assault reports

Research Paper

4 April 2023

Academic support for *'True or false, or somewhere between?'*

"I appreciate the care and meticulous attention to detail shown by Tom Nankivell in this manuscript. … Tom makes a valuable contribution to the field with this manuscript, enhancing the credibility of research on false reporting with a methodical, transparent, and eyes-wide-open assessment of this most complex and challenging issue."

Dr Kim Lonsway
Director of Research
End Violence Against Women International
(Lead researcher for the 2009 US high-quality prevalence study on false sexual assault reports)

"This is an important and overdue study. Given the paucity of research on false allegations of sexual offences, but also the valid concerns about giving disproportionate attention to them, this is just what is needed: a quantitative analysis which raises pertinent issues regarding estimations of prevalence, definitions, and methodologies, but without, it seems, taking any side or being polemical."

Dr Ros Burnett
Centre for Criminology, University of Oxford
(Editor of 'Wrongful allegations of sexual and child abuse', OUP 2016)

About Gander Research

Gander Research aims to produce rigorous research on selected gender issues. We have extensive experience in public policy research and analysis. We intend to interact with other researchers, make submissions to relevant government bodies, and publish our work. We also encourage others to collaborate with us and use our research. To read more Gander Research, provide feedback or join our mailing list, visit www.ganderresearch.org.

About this study

This research paper was compiled by Tom Nankivell and John Papadimitriou with input from several former colleagues. We sought clarifications and feedback from a number of sexual violence researchers, including authors of all the recent (post-2000) high-quality prevalence studies, at various points during the paper's gestation. We also circulated a full draft to them in December 2022, with a further invitation for comments, corrections or suggestions. We thank those academics who responded, and welcome further feedback.

True or false, or somewhere between?

A review of the high-quality studies on the prevalence of false sexual assault reports

Tom Nankivell and John Papadimitriou[1]

KP

Key points

Prominent academics in the sexual violence field agree that the prevalence of false sexual assault allegations made to police is estimable, known within tight bounds, and very low. The most commonly cited estimates are that "approximately 5%" or "between 2% and 10%" of reports are false.

The studies on which the consensus is based have several limitations. Most critically, the rules they follow to determine whether to classify a report as "false" exclude many false and potentially false reports. Their estimates are premised on there being no false allegations among the many equivocal or ambiguous cases classified as having insufficient evidence, or where the alleged victim withdrew their complaint, or where the accused was tried but acquitted. Some of the studies also suffer from incomplete or poor-quality data, limited interview response rates and mathematical shortcomings.

The studies' prevalence rates are better viewed as lower bound estimates. Using more realistic but still modest assumptions about the share of false reports in other categories, we show that the actual prevalence rate could be materially greater than the studies' estimates. However, the inherent difficulties in separating fact from fiction in many sexual assault cases precludes pinpointing the actual rate, or devising meaningful upper bound estimates.

Our findings have implications for how the sexual violence literature classifies false reports and how it describes and communicates prevalence rate estimates.

Introduction

Rape and other forms of sexual assault can ruin lives. Most perpetrators are male; the victims are mainly female. There is little doubt that the majority of sexual assault allegations made to police are true.[2]

However, the exact prevalence of false reports of sexual assault has long been disputed. Published estimates of the prevalence rate have varied wildly, from under 2% to upwards of 40% — with the odd outlier much higher again (Rumney 2006, 137). This variation stems from differences in how studies define false reports (if they do), their samples and methodologies, and the inherent challenges in separating fact from fiction in many sexual assault cases.

The contemporary academic consensus is that false reports are rare

Nevertheless, a consensus has emerged among prominent academics in the sexual violence field that the prevalence rate is estimable, known within tight bounds, and very low.[3] This is based principally on the results of studies that use what these academics regard as sophisticated, rigorous and reliable methodologies. In these "high-quality" studies, as we shall call them, typically the researchers independently analyse crime reports or summaries, use accepted or "correct" definitions of false reports, and attempt to corroborate information with victims, police officers and other trained individuals (Weiser 2017, 51).

Referring to a selection of high-quality studies, Lisak, Gardinier, Nicksa & Cotely (2010) stated that the credible prevalence rate estimates lay between 2% and 10%. A meta-analysis by Ferguson & Malouff (2016) of a similar selection of studies (box 1) generated a weighted prevalence rate estimate of 5%.

1

Delineating the high-quality studies

Although Rumney (2006, Table 1) listed 20 prevalence studies conducted between 1974 and 2005, contemporary scholars have been critical of most of these (see, for example, Lisak, Gardinier, Nicksa & Cotely 2010, 1319-1322). Concerns include that the studies often lack transparency and rigour and/or accept police classifications at face value notwithstanding that police may hold biases against classes of victims and are known to misclassify or inconsistently classify cases.

The criteria used by Lisak et al. (2010, 1324-1327) for delineating credible studies included that the studies clearly define what constitutes a false report, clearly explain the source of the data used, and make some effort to evaluate the data they receive from law enforcement agencies. They included the following studies:

- 1977 Toronto study by Clark & Lewis
- 1979 Philadelphia study by McCahill, Meyer & Fischman
- 1992 British study by Grace, Lloyd & Smith
- 1999 British study by Harris & Grace[4]
- 2005 British study by Kelly, Lovett & Regan
- 2006 Australian study by Heenan & Murray
- 2009 US cities study by Lonsway & Archambault
- 2010 US university study by Lisak, Gardinier, Nicksa & Cotely.

For their meta-analysis, Ferguson & Malouff (2016, 1187) searched academic databases, review articles and published empirical reports to identify candidates for inclusion. Compared to the Lisak et al. list and using similar criteria, they excluded the two earlier British studies while adding the later 2014 Los Angeles study by Spohn, White & Tellis.

There do not appear to have been other studies that would meet the criteria published since then.[5]

Issues with the way the studies classify reports as false

However, Ferguson & Malouff (2016, 1187-89) cautioned that the prevalence estimates from the high-quality studies are constrained by the methods and classifications those studies adopt. The studies follow official case classification/counting rules used by police when determining whether to classify a report as false. The International Association of Chiefs of Police (IACP 2005, 12-13) articulated the approach as follows:

> The determination that a report of sexual assault is false can be made only if the evidence establishes that no crime was committed or attempted. This determination can be made only after a thorough investigation. This should not be confused with an investigation that fails to prove a sexual assault occurred. In that case, the investigation would be labelled unsubstantiated.

There are two aspects to this statement: a *definition* and a *classification rule*. The *definition* of a false report embodied in the statement is "a report of a crime that was not committed or attempted". (This has been explicitly recognised as the most reasonable definition of a false report by key sexual violence researchers[6], and is used in this paper too). In addition, the statement explains when police can classify a false report as such. That *classification rule* is: only after a thorough investigation that yields evidence that proves or establishes that the alleged criminal act did not take place.

Crucially, this does not mean that false reports that have not (yet) been determined to be false are not false; just that police should not record them in their "false" category. Rather, to follow their rules, police should record such reports in other categories, such as "case ongoing", "no further action" due, for example, to insufficient evidence, or "complaint withdrawn". (The precise categories and labels used can vary between jurisdictions and across time).

These classification rules may well be reasonable for official policing purposes, helping to signal whether and why an investigation is closed or still active, and because making a false report can be a chargeable offence, which is a relevant consideration for police forces.

However, the IACP's false report classification rule is problematic for the very different purposes for which the high-quality studies' estimates are used, such as to draw inferences about the share of sexual assault reports that are valid and about the general credibility that should attach to complainants and alleged perpetrators. Adhering to the IACP's classification rule leads researchers to make no allowance in their estimates for those false reports for which the alleged offence cannot be, or has not been, disproven to the required standard. Given that sexual assaults often have few if any witnesses, scant or ambiguous evidence and can come down to "her word versus his", following this approach makes it possible that many false reports will not be counted as such.

Other bases on which the high-quality studies may exclude reports from their prevalence estimates are also noteworthy. For example, researchers might not classify an allegation of sexual assault as false when the complainant was affected by alcohol, drugs, or poor mental health; lacked awareness of the law (and so alleged an offence when the actions complained of were legal); or mistook the identity of the offender.[7] Some of the studies also exclude false reports made by third parties such as family or friends of the alleged victim. As we shall see, some of these approaches are not congruent with the IACP *definition* of a false report, mentioned earlier. The studies also typically look at only those false allegations made to police.

Ferguson & Malouff (2016, 1187) explain the reasons for, and ramifications of, the researchers classifying sexual assault allegations in this way:

> Although limiting the sample, [adopting these constraints is] a necessary step as it prevents opening the floodgates to many equivocal cases that are suspected but not demonstrably false. It errs on the side of caution by not including cases in doubt, mistaken cases, or those claims made to anyone other than police. Use of such a conservative definition is not meant to imply that all other cases are true reports, but just that they cannot responsibly be deemed confirmed false.[8]

In effect, Ferguson & Malouff imply that the estimates might better be thought of as "lower bound" estimates of the rate of false allegations reported to police. They hint that "upper bound" estimates might be derived by adding "equivocal" or "suspicious" cases to them.

About this paper

Against this backdrop, we reviewed the studies analysed by Ferguson & Malouff to see whether it is possible to determine a credible upper bound prevalence estimate for their samples.

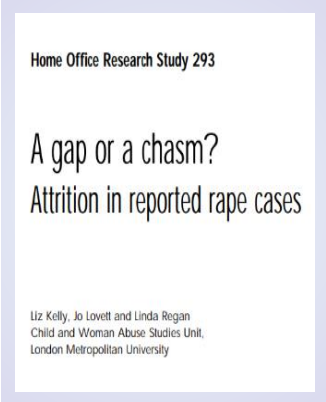
For each of the studies, we describe the study's background, outline the sample and methodology, report its (lower bound) prevalence estimate, explain why that estimate does not cover the full field of false reports in the study's sample, and explore the scope for devising an upper bound estimate from the information in the study.

We start with the major and influential 2005 British Home Office study by Kelly, Lovett & Regan, which we have examined in the most detail. Our focus is on the potential for false cases among the very large number of cases in other categories, such as "insufficient evidence", "victim withdrawal" and "acquittals", that the study in effect assumes contain no false reports. With the nature and a sense of the potential magnitude of this problem established, we can address the equivalent problems in the other studies more swiftly.

Our main conclusion is that the high-quality studies have some critical shortcomings; and that the actual prevalence rate is not and cannot be known but is potentially materially greater than the studies' estimates indicate. In turn, we suggest some implications for how the sexual violence literature classifies false reports and how it describes and communicates prevalence rate estimates.

The 2005 British study

To help understand why only a small share of sexual assault accusations translate into convictions, Kelly, Lovett & Regan (2005) prospectively tracked three-and-a-half thousand rape cases to and through the UK criminal justice system. The sample was drawn from complaints made to a selection of sexual assault referral centres and police sites in inner city, suburban and rural areas in Britain during 2000-2002.



The authors analysed data and qualitative information on the cases and tried to drill down on the attitudes and practices of police and others in the system. Data were collected on the alleged assaults, the complainants, the alleged perpetrators, forensic exams, services accessed by complainants, and the criminal justice process and outcomes, where known.

Victims opted into a series of questionnaires and interviews for a subset of 228 of these cases. Where possible, the researchers examined related police statements and forensics reports, and interviewed police and other experts.

The researchers also circulated questionnaires to police who investigate sexual assaults, although the authors acknowledged that, among other methodological challenges they encountered, "[t]he data on the proformas limit the extent to which one can assess the police designations" (Kelly et al. 2005, 50).

In looking at the early causes of complaint attrition, the authors estimated that just 2.5–3% of rape reports made to police were false, much lower than the police's own estimates.

How did Kelly, Lovett & Regan estimate the prevalence rate?

The study started with a sample of 3527 complaints, some made directly to police, and others made to sexual assault referral centres. However, a number of the complaints reported to the referral centres did not translate into reports to police. And of those where reports were made to police, in some cases the researchers received insufficient information to know or analyse the complaints and/or their outcomes (including because police sometimes failed to complete proformas or the cases had not been concluded). The upshot was that the researchers were left with a sample of 2643 complaints reported to police, and of 2284 complaints for which the outcomes and/or research categories were known (Kelly et al. 2005, 38-39).

The study examined how many of these could be deemed "false" using police internal counting rules which specified that "this category should be limited to cases where either there is a clear and credible admission by the complainants, or where there are strong evidential grounds" (Kelly et al. 2005, 50). The process used was explained on pages 47-50 of the study:

- Some 216 of the complaints were classified as false by police.
- Of these, the study authors had access to sufficient information to make their own assessments in 144 cases.
- The authors classified the falsity of these cases as either "probable" (primarily where complainants accounts were referred to), "possible" (mainly where there was some evidential basis for the false report classification), or "uncertain" (including where victim characteristics were seen to make them less believable). 44, 33 and 77 cases were identified in each of these categories, respectively.
- The authors excluded the 77 "uncertain" cases from their own estimates of the prevalence rate, counting only the "probable" and "possible" cases to arrive at a total number of false cases of 67.

The authors concluded:

> If the proportion of false complaints on the basis of the probable and possible cases are recalculated, rates of 3% are obtained, both for all reported cases (n=67 of 2643), and for those where the outcome is known (n=67 of 2284).[9]

Some miscalculations

There are two mathematical shortcomings in the study's prevalence rate calculations.[10]

The authors did not adjust their estimates of the total number of cases they deemed false (which is the numerator in a prevalence rate estimate) to the same basis as that used for the total number of cases (the denominator). The authors divided the 67 cases they deemed false by "2284", which is the total number of cases reported to police with known outcomes and/or research categories. However, the 67 false cases are just a subset of the 144 cases for which the authors had sufficient information to make their own assessments. Those 144 cases are themselves a subset of the 216 cases that were classed as false by police. Accordingly, the authors needed to inflate their "67" false case figure by 216/144 to be comparing false cases with total cases on like-with-like terms. This means that the prevalence rate calculation should be (67x216/144)/2284, which yields a false report prevalence rate of 4.4%.[11]

Kelly and her colleagues also seem to have misreported or mistallied some case count numbers, which may have caused them to understate the prevalence rate slightly (box 2). In the worst case, adjusting the prevalence rate estimate for this error would lift the (corrected) estimate from 4.4% to 5.1%.[12]

2

Miscounted cases

To summarise the study's prevalence rate calculations, Kelly et al. (2005, 50) said that out of a total of 144 cases for which there was information, they added the 44 probable cases to the 33 possible cases to arrive at a sum of probables and possibles of 67, with 77 cases remaining uncertain.

This does not add up correctly, because 33+44 = 77, not 67; and 44+33+77 = 154, not 144. There are three main ways this error of 10 could have arisen:

1) the "44" figure or the "33" figure are incorrect numbers (eg one should be 10 less than it is)
2) the "44", "33" and "144" are all correct numbers, but the "67" figure (for probables plus possibles) should be "77" — and the "77" figure for uncertain cases should actually be "67"
3) the "44", "33" and "77" (for uncertain cases) are all correct numbers, but the "67" figure (for probables plus possibles) should be "77", and the total cases for which there is information "154" rather than "144".

Of these three error pathways, the first would have no impact on the calculated prevalence rate. However, if the error occurred via the second or the third pathway, the prevalence rate would be higher than the study's calculations suggest.

False reports excluded from the study's prevalence estimates

The more problematic aspect of the study is that, consistent with the broad approach in other high-quality prevalence studies, Kelly, Lovett & Regan adhered to internal police case counting rules and so followed the IACP approach for classifying reports as false. As noted earlier, together with other conventions adopted by the studies, this way of classifying false reports constrains the estimated prevalence rate.

To gain a more realistic sense of the prevalence of false reports, it is necessary to explore the scope for false reports in the study's sample over and above the study's estimates.

For 2284 cases in their sample, Kelly and her colleagues traced progress through the criminal justice system and also classified the reasons for cases terminating when they did (table A). The study authors grouped these reasons into six "attrition points".

In the following sub-sections, we highlight the categories in these different attrition points that are most likely to harbour at least some false allegations, additional to those covered in the study's lower bound estimate. The main candidates are:

- no evidence of assault and "uncertain" false reports
- insufficient evidence etc
- withdrawn complaints
- acquittals.

For the more than 1500 cases in these categories, the study authors effectively assumed that *none* was false. While it is not possible to know the real number of false cases in the categories, for illustrative purposes we set out what some more realistic assumptions about the share of false cases would imply for the study's overall prevalence rate estimate.

A

Reasons for attrition in 2284 rape cases

As categorised by Kelly et al. (2005, 40) for cases where outcomes and/or research categories known

Case phase/reason for attrition	No. of cases	% of total cases
Police phase	**1817**	**79.6**
Insufficient evidence	386	16.9
Victim withdrew	318	13.9
Victim did not complete initial process	315	13.8
Offender not identified	239	10.5
False allegation	216	9.5
No evidence of assault	83	3.6
No prospect of conviction	37	1.6
Not in public interest	20	0.9
Other	67	2.9
Reason unknown	136	6.0
Criminal Prosecution Service phase	**145**	**6.3**
Caution/final reprimand	9	0.4
Discontinued	38	1.7
Victim withdrew	25	1.1
Suspect fled	1	<.1
Pending trial	72	3.2
Trial phase	**322**	**14.1**
Suspect fled	1	<.1
Victim withdrew	15	0.7
Discontinued/withdrawn at court	19	0.8
Acquittal	104	4.6
Guilty plea	89	3.9
Guilty plea/conviction unclear	17	0.7
Partial conviction	11	0.5
Conviction	66	2.9
Total	**2284**	**100**

Attrition point 2 – No evidence of assault and false reports

Almost 300 cases that advanced no further than the police stage were closed due to what was classed as "no evidence of assault" or "false allegations". The study authors discussed these groups together, as "attrition point 2", in pages 46-53 of the study. (The first attrition point was where people who had reported to a sexual assault referral centre did not go on to report the allegation to police. The reasons for this are not relevant for determining the share of reports to police that are false).

Of the 83 cases categorised as "no evidence of assault", the authors had sufficient information to analyse 56. From these, at least 11 are false allegations under the ordinary meaning of that term. These complaints were made by family members or someone other than the alleged victim and proved to be wrong. There were an additional 31 cases were where the complainant had lost consciousness or was affected by drugs during a possible incident, but subsequent forensics and toxicology indicated that no sex had occurred. Of these cases, Kelly et al. (2005, 47) stated that it is "an open question whether this group of cases should be combined with false allegations in police record-keeping", and did not include them in the numerator of the study's prevalence estimate. In our view, Kelly's approach is problematic (see box 3.)

How should we deal with the "open question" cases?

Neither of the reasons Kelly et al. (2005, 47) gave for excluding the 31 cases went to the question of whether the reports were actually false. Those reasons were: "firstly 'no evidence of assault' is not a contested category; and secondly, given the data presented later on how 'previous allegations' undermine the credibility of complainants, there is a danger that such records might in the future be read, erroneously, in this way." However, in a later article, Kelly (2010, 1349) argued for their exclusion on the basis that, in most cases, the approach to police did not really constitute an "allegation" but was instead a question or concern answered relatively easily.

We note that Kelly argued this in relation to only "most" of the 31 cases, which presumably leaves grounds for including the remaining, unspecified share as false reports. However, to the extent that any or all of the cases in this category are not included in the numerator of the study's prevalence estimate, on the basis that they were not really allegations of sexual assault, the corollary is that they should be removed from the denominator (along with other cases in the study's "no evidence of assault" category).[13]

As discussed earlier, in only 216 cases did the police classify a complaint as false. In devising their prevalence estimate, Kelly and her colleagues excluded around half the cases on which they had information arguing that the police pro-formas indicated that they had made their decisions based mainly on victim characteristics, in contravention of the police case classification rules. This is a reasonable approach given the study's adherence to those rules. However, it does not follow that a good number of these "uncertain" cases were not false in the common sense of the term or in accord with the IACP definition (as distinct from the IACP classification rule).

If just one third of the 77 "uncertain" cases were counted together with the 11 false allegations by family members etc, this would raise the prevalence rate for the study's sample by nearly 2.5 percentage points.[14] Adding more of the "uncertain" cases and/or a share of the 31 "open question" cases would obviously lead to a greater increase.

Attrition point 3 – Insufficient evidence etc

Some 386 cases that advanced no further than the police stage were closed due to "insufficient evidence". These cases are discussed in pages 53—58 of the study, together with the 239 cases in which an offender was not identified and the 37 cases for which police deemed there was "no prospect of conviction", as "attrition point 3 — insufficient evidence".

The study authors reported a range of reasons why police had discontinued these cases, including:

- the complainant's inability to recount clearly what happened (in some cases because the complainant had severe learning difficulties, mental health problems, or had been impaired by drugs or alcohol)
- the failure to identify or trace an alleged assailant
- assessments by the police (often in consultation with the Crown Prosecution Service (CPS)) that cases are weak or stand little prospect. (Some of the police reasons given for this were that a complainant had been unreliable, there was no forensic or CCTV evidence, or sex evidently occurred but lack of consent was unclear or disputed).

There are good reasons why, particularly following rape, some victims may be unable to provide a clear account, and also why the available evidence may be limited. The study also highlighted some questionable bases for police judgments about the veracity of a number of the complainants, and some shoddy investigatory practices that may have allowed evidence to be missed or assailants to evade detection. None of these matters mean that a particular complaint is false, and there is no reason to doubt that many were true.

Equally though, it cannot be ruled out that, for at least some of the rape allegations in this category, the lack of sufficient evidence, an identifiable perpetrator and/or a good prospect of conviction partly reflects that the allegations were in fact false. It is not possible from the information in the study to judge the share of false or potentially false allegations in the category. However, for each tenth of these 662 cases that were counted as false reports, the study's prevalence rate would increase by 3 percentage points.[15]

Attrition point 4 – Early victim withdrawal

The alleged victim withdrew their complaint, or declined to complete the initial police process, in 633 cases that did not advance beyond the investigatory phase. Pages 59-69 of the study cover these cases. (There were also 40 withdrawals after cases were referred to the CPS and when set for trial.)

There is little doubt that the majority of these reports would have been true, even though the complainant withdrew at various points in the process. The study sets out a range of reasons why victims might withdraw, including:

- to avoid the awkwardness of being interviewed or forensically examined, particularly by males, in the immediate aftermath of a rape
- to avoid the disruption and embarrassment that a police investigation might cause (for example, where friends and associates are questioned about the complainant's lifestyle/behaviour)
- a lack or loss of faith in the police or the justice system, sometimes following non-supportive interactions with police or prosecutors, harsh or sceptical questioning, or the police's failure to identify a suspect
- a fear of the court process, including giving public testimony about intimate matters, and having one's behaviour cross-examined and judged, with the risk that an acquittal may follow
- a decision to "protect" the perpetrator, where a family member or friend, or due to pressure, harassment, or intimidation from the perpetrator or their associates
- a desire to "put the event behind them" and move on with life.

These are all understandable reasons for withdrawing that do not necessarily speak to the truth or falsity of the original allegation.

Even so, at least some reports in this category are likely to have been false, with some complainants withdrawing because they had had second thoughts about the incident. Supporting this, in nearly one-quarter (15/66) of the cases where there was some information on why the complainants withdrew early, the study authors said that this was due to the complainant being uncertain about whether they were raped, including in some cases thinking that the sex may have been consensual (Kelly et al. 2005, 64). It is also possible that some complainants simply fabricated their allegation but then developed cold feet, perhaps due to ethical qualms or an increasing fear of exposure.

Again, we cannot know what share of all complaints in this category was true and what share was false. However, for each tenth of these 633 cases that are counted as false reports, the study's prevalence rate would increase by nearly 3 percentage points.[16]

Attrition point 6 – Acquittals

322 cases made it to the trial phase (attrition point 6, discussed on page 71-77 of the study). Of these, the majority resulted in guilty pleas and/or convictions but there were also 104 acquittals (plus a small number of cases of complainant withdrawal, discussed earlier).

The "acquittals" category is another where false cases may lurk. Of course, given the need to prove a case beyond reasonable doubt to secure a criminal conviction, and the unclear or disputed facts that often surround sexual assault allegations, an acquittal is not proof of factual innocence. Nevertheless, where evidence including complainant testimony is laid before a court, it is reasonable to presume that, in a share of these cases, the court's decision to acquit reflects a view that the allegation brought forward was false or likely false.

For each tenth of these 104 cases that were counted as false reports, the study's prevalence rate would increase by 0.4 percentage points.

False cases in other categories?

There could also be some false cases in other categories, such as among the 40 complainants who withdrew at the CPS or trial phase, or the 57 cases where the CPS discontinued the case or where charges were withdrawn during trials. There were also more than 200 cases discontinued at the police stage for which the reason was listed as "unknown" or "other", and around 70 cases at the CPS stage "pending trial" for which the outcome — withdrawal, conviction or acquittal — had not at that time been determined (table A).

Can we pinpoint the prevalence rate or devise a meaningful range?

Kelly et al.'s (2005) prevalence rate estimate was based on the extreme assumption that none of the 2284 cases outside those labelled false under police counting rules were false.

The calculations in this paper show that, even with reasonably modest assumptions about the actual level of false allegations in other categories, the prevalence rate for the study's sample would easily exceed 10% and could approach 15%[17] — that is, 4 to 6 times higher than the widely quoted figure for the study of 2.5%. Less conservative assumptions would generate estimates above that level.

However, the information in the study provides no rigorous way to pinpoint the "correct" or actual prevalence rates for these other categories, and thus the actual overall prevalence rate.

This raises the question of whether there is some way to specify an "upper bound" estimate for the prevalence rate that would provide a credible and meaningful range within which the actual prevalence rate statistic would sit?

An obvious approach would be to adopt the inverse of the conviction rate, on the assumption that none of the cases eliciting a guilty plea and/or conviction were false. There have of course been some well-publicised cases of wrongful conviction, but these are presumably very rare. However, with only 8% of cases ending in conviction, this would yield an upper bound estimate of 92%. Given that there are sound reasons to believe that the majority of rape allegations are true, the 92% figure is too high, and would leave too wide a range from lower bound to upper bound, to be useful or very meaningful.

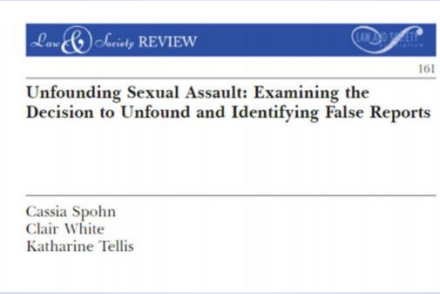
Using the inverse of the rates of cases making it to the trial phase, or to the CPS phase (which would yield prevalence rates of 86% and 80%, respectively), would not overcome the problem of generating too wide a range to be useful. And they would be less credible as "upper bound" estimates in the sense that it can be safely contended that a number of the cases reaching those phases are in fact false.

The upshot is that it is not possible to devise a credible and meaningful upper bound estimate for the study.

The 2014 Los Angeles study

Spohn, White & Tellis (2014) examined sexual assault cases in Los Angeles from 2008 involving female complainants over the age of 12. The LAPD categorised such cases as either:

- cleared by arrest
- cleared by exceptional means
- ongoing
- unfounded.


Law & Society REVIEW

161

Unfounding Sexual Assault: Examining the Decision to Unfound and Identifying False Reports

Cassia Spohn
Clair White
Katharine Tellis

The study's aim was to understand unfounding decisions and to provide a better estimate of the prevalence of false reports than earlier studies.

From a starting stratified sample of 401 cases, the study team scrutinised the case files of 81 complaints the LAPD had classed as "unfounded", to determine how many were "false".

The three co-authors reviewed each of the cases and cross-checked their judgments. The case files were detailed and included crime reports by the relevant patrol officers and detectives; accounts of statements made by complainants and, where taken or available, by alleged perpetrators and any witnesses; and descriptions of crime scene evidence and results of medical exams.[18] The study team also interviewed LAPD detectives experienced in sexual assault cases.

The study classified false cases in line with the IACP approach. A false report was classified as such only if the case file contained evidence that the police had conducted a thorough investigation that led them to conclude that the complainant deliberately fabricated the allegation. False allegations for which there was no evidence that the complainant intentionally lied about the incident were classified as "baseless" rather than "false".

Of the 81 unfounded cases, the authors categorised 55 as false. In the majority of these, the complainant recanted her allegation and there was evidence that a sexual assault did not occur. However, the authors did not assume that complainants who recanted had necessarily filed a false report.[19] The authors calculated that the 55 cases categorised as false translated to a prevalence rate of 4.5%.

Among the 26 other unfounded reports, the authors classed 5 as baseless and 11 as unclear as to whether they were false/baseless or not. (Note that there is a general argument for either counting baseless reports together with false reports in the numerator of prevalence estimates, or for removing them from the denominator).[20] The other 10 were mainly cases where the complainant recanted but there was evidence that this was motivated by fear or pressure, or a lack of interest in proceeding with the case. To the extent that some of these 26 cases were indeed false reports, this would cause a small increase in the prevalence rate for the sample.

However, the larger gap in the study is that it did not account for the scope for false reports outside these categories, such as in the "ongoing" category. In effect, by classifying false allegations in line with the IACP approach and examining only those allegations the police had "unfounded", the authors were deriving estimates that incorporated an assumption that none of the cases in the other categories were false. Yet as per our discussion of the comparable categories in the 2005 British study, while it is feasible that the bulk of the allegations in the other categories were of real sexual assaults, it is also reasonable to assume that a subset were false. Indeed, Spohn, White & Tellis (2014, 186) acknowledged that:

> … our interviews with LAPD detectives revealed that some of them were reluctant to categorize a case as "unfounded", even if they believed that it was false or baseless; these detectives reported that they would clear the case by exceptional means or keep the case open. In addition, we have no way of knowing if there were false allegations that were not recognized as such and that were cleared by arrest or exceptional means. Considered together, these data limitations suggest that the rate of false reports among rapes reported to the LAPD in 2008 may be somewhat higher than 4.5%.

The study's design did not allow the question of "how much higher?" to be gauged. Thus, while the authors believed they had overcome several deficiencies associated with earlier research on the topic, it is not possible from their study to devise a robust prevalence estimate, or even a reasonable lower-bound/upper-bound range, for false sexual assault allegations in Los Angeles in 2008.

The 2010 US university study

Lisak, Gardinier, Nicksa & Cotely (2010) examined 136 sexual assault allegations made to a north-eastern US university police department between 1998 and 2007. The authors were granted access to the case summaries which included a chronology of the investigation, information from interviews with the alleged victims, perpetrators and any witnesses, and other evidence collected. The authors also interviewed senior police with knowledge of the cases.

Symposium on False Allegations of Rape

False Allegations of Sexual Assault: An Analysis of Ten Years of Reported Cases

David Lisak[1], Lori Gardinier[2], Sarah C. Nicksa[2], and Ashley M. Cote[2]

Abstract

One of the most controversial disputes affecting the discourse related to violence against women is the dispute about the frequency of false allegations of sexual assault. In an effort to add clarity to the discourse, published research on false allegations is critiqued, and the results of a new study described. All cases (N = 136) of sexual assault reported to a major Northeastern university over a 10-year period are analyzed to determine the percentage of false allegations. Of the 136 cases of sexual assault reported over the 10-year period, 8 (5.9%) are coded as false allegations. These results, taken in the context of an examination of previous research, indicate that the prevalence of false allegations is between 2% and 10%.

Keywords

false allegations, sexual assault

Rape is unique. No other violent crime is so fraught with controversy, so enmeshed in dispute and in the politics of gender and sexuality. For example, despite decades of careful social science research, prevalence rates are still frequently challenged on political grounds, and bold assertions are made in the absence of any data (e.g., MacDonald, 2008; Roiphe, 1993). And within the domain of rape, the most highly charged area of debate concerns the issue of false allegations. For centuries, it has been asserted and assumed that women "cry rape," that a large proportion of rape allegations are maliciously concocted for purposes of revenge or other motives. Most famously, Sir Matthew Hale, a chief justice of the court of the King's bench of England, expressed this view in a form that became the basis for

[1]University of Massachusetts Boston
[2]Northeastern University, Boston, MA

Corresponding Author:
David Lisak, Department of Psychology, University of Massachusetts Boston, 100 Morrissey Blvd, Boston MA, 02125
Email: dlisak5@gmail.com

The researchers assigned allegations to the following categories:

- false
- did not proceed to prosecution or disciplinary action
- proceeded to prosecution, disciplinary action or other administrative action
- insufficient information to assign.

The classification of a report as false followed IACP guidelines and required that a thorough investigation was pursued which had yielded evidence that the reported sexual assault had not occurred. The study authors noted that "if key elements of a victim's account were internally inconsistent and directly contradicted by multiple witnesses and if the victim then altered those key elements of his or her account, investigators might conclude that the report was false." (Lisak et al. 2010, 1328). That conclusion would have been based not on a single interview, or on intuitions about the credibility of the victim, but on a preponderance of evidence gathered over the course of a thorough investigation (where a "thorough investigation" would potentially require multiple interviews of the accused, the complainant, and other witnesses, and where applicable, the collection of other forensic evidence (e.g., medical records, security camera records).

Under this approach, the authors found that just 8 allegations were confirmed to be false, resulting in a 5.9% prevalence rate estimate (ie 8/136). The authors noted that "[t]hese results are consistent with those of other studies that have used similar methodologies to determine the prevalence of false rape reporting". (ibid. 1329). (However, it can be argued that the authors should have reduced the denominator by 19 cases — the number in the fourth category termed "insufficient information to assign" — which would yield an adjusted (lower bound) estimate of 6.8% (ie 8/117)).[21]

What is not discussed in the study is what proportion of cases in the other categories were suspicious or equivocal and thus could also be false. In effect, by classifying false allegations in line with the IACP guidelines, the estimates are based on the assumption that none of these cases were false.

Some 61 cases were assigned to the second category: "case did not proceed". The authors explained that this classification was applied if there was insufficient evidence, the complainant withdrew from the process or could not identify a perpetrator, or if the incident did not meet the legal elements of the crime of sexual assault. As per our earlier discussion of the comparable categories in the 2005 British study, while it is feasible that the bulk of these allegations were of real sexual assaults, it is also reasonable to assume that a subset were false allegations. For each tenth of these allegations that were false, the prevalence rate for the study's sample would be 4.5% higher than reported. So, for example:

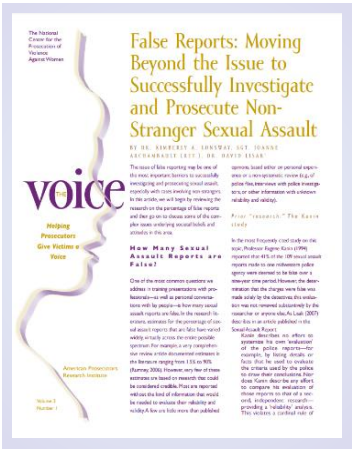
- if the frequency of false cases in the second category was one-in-ten, the overall prevalence rate would rise to 10.4% (or 12.1% if the unassigned cases were removed from the denominator)
- if one-in-five of these cases were false, the prevalence rate would rise to 14.9% (or 17.3% with the denominator adjusted).

Higher rates of false reports in this category would of course translate into an even higher overall rate.

The third category contained a further 48 cases: those that proceeded for prosecution or disciplinary or administrative action. Most of these cases can be assumed to have been for real sexual assaults, with few if any false. However, for each tenth of these cases that were false, the overall prevalence rate would rise by a further 3.5% (or 4.1% with the denominator adjusted).

While it is possible that the prevalence rate of false reports could thus be much higher than the published figure, the study does not provide a basis for determining the actual rate or even a reasonable upper bound estimate.[22]

The 2009 US cities study


False Reports: Moving Beyond the Issue to Successfully Investigate and Prosecute Non-Stranger Sexual Assault

voice

Helping Prosecutors Give Victims a Voice

How Many Sexual Assault Reports are False?

Lonsway, Archambault & Lisak (2009) reported on a multi-year project by End Violence Against Women International (EVAWI) in a diverse sample of eight US cities. The project was intended, in part, to "set new national standards for effectively prosecuting sexual assault".

Each of the cities had an established Sexual Assault Response Team (SART), with representatives from law enforcement, prosecution, health care, and victim advocacy. SART members agreed on the definitions and approaches adopted in the study. Along with numerous other professionals in each community, they received training and ongoing technical assistance from EVAWI on topics related to sexual assault response and investigation.

Altogether some 40 agencies were involved, which were asked to provide data on all felony-level sexual assault cases handled during an 18–24 month data collection period between 2004 and 2006. Cases were excluded if the alleged victim was unable to consent due to age, professional relationship or institutionalisation.

The study gathered data from almost 9400 sexual assault cases from the various agencies involved, including 2059 police reports from which the study's estimate of false reports was derived.

Information on police case dispositions was available for 1984 of these 2059 reports, with the police classifying them as one of the following:

- closed as informational reports (17.9%)
- unfounded: false (7.1%)
- unfounded: baseless (8.5%)
- suspended/inactivated (28.6%)
- exceptionally cleared (17.9%)
- cleared by arrest (20.0%).[23]

A feature of the study was the care taken to ensure that reports were classified as false in accordance with the official crime reporting guidelines. Thus, "all participating law enforcement agencies were provided training and technical assistance in an ongoing way to ensure that they were applying consistent definitions of a false report. In addition, a random sample of cases was checked for data entry errors" (Lonsway, Archambault & Lisak 2009, 2). Participating law enforcement agencies were advised that a report was only to be classified false on the basis of evidence that a crime was not committed nor was attempted. Thus, a case should not have been classified as false if the investigation produced "insufficient evidence" or if no investigation was conducted.

Against this backdrop, of the 1984 cases that police classified, 140 (or 7.1%) were classified as false.

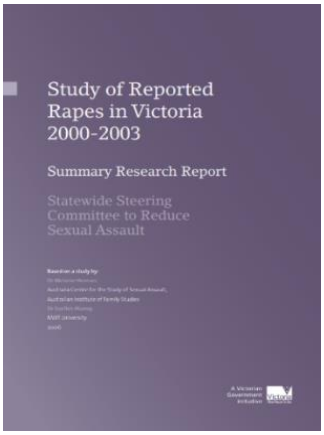
One issue with this estimate is that cases recorded as informational reports are counted in its denominator, yet these reports are where the legal elements necessary to establish a crime of sexual assault have not been deemed present. For example, a woman might report a situation where she felt pressured into sexual contact, but the coercion would not meet the criteria for a forcible sexual assault. By classifying them as informational reports, they are, by definition, not crime reports. There is therefore a case for reducing the denominator by 17.9% (or at least by a share of this figure[24]), which in turn would increase the estimated prevalence rate for the study from 7.1% to 8.6%[25] (or 9.6% if the baseless reports were similarly excluded).

Without further investigation of the other cases included in the project's sample, there was no way to explore whether any additional reports (classified in other categories) may have also been false reports.[26]

Nevertheless, consistent with equivalent categories in the other studies, it is possible that there were some false reports among those other cases. This seems most likely for cases that were suspended/ deactivated or in other ways terminated due to lack of evidence to support an arrest. This could be because the victim recanted or withdrew participation in the investigation, provided false statements or did not provide a statement etc. These reasons do not necessarily or of themselves mean that the reports were false, but it is plausible that some were. There may also have been a (presumably smaller) number of false cases among those "cleared by arrest", given that arrests only required that police deemed that cases met the probable cause standard of proof and that a number of such cases did not result in findings of guilt. On the other hand, it seems possible that there would not have been many false cases that were misclassified as "cleared by exceptional means" in this study.[27]

Again, however, it is not possible to determine a meaningful upper bound estimate of false reports for the study's sample, or the actual rate of false reports.

The 2006 Australian study



In a study for the Victorian Office of Women's Policy, Heenan & Murray (2006) looked at a sample of 850 rape allegations reported to Victoria Police between 2000-2003. As well as describing the characteristics of the cases, the study sought to examine influences on the outcomes of rape investigations.

The sample of cases did not exclude male or young victims and was weighted to ensure representation of crime reports across metropolitan and regional Victoria.

The study authors had access to case information in the form of data contained on the Victoria Police Law Enforcement Assistance Program (LEAP) database. It provided the capacity for police to record information about the alleged victim and offender, the nature of the offence, and the nature, scope and outcome of investigations and action taken. However, the level of information in the LEAP database was sometimes limited, and the researchers ended with an incomplete data set for 341 of the cases examined in the sample (Murray & Heenan 2012, 359). Data problems hampered the study's ability to review police attitudes and perceptions. There were also 38 missing cases.

The study found that police classed just 17 of the 812 cases for which records were available as "false reports". The resultant prevalence rate of 2.1% made its way into the Ferguson & Malouff meta-analysis, as one of the lower prevalence estimates.

However, Heenan & Murray (2006, 20) effectively acknowledged that the 2.1% figure was a lower bound estimate:

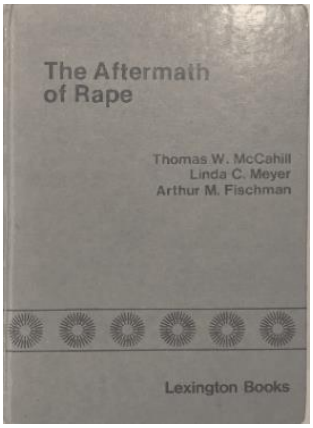
> In 17 cases (2.1%), the case outcome was clearly categorised as a false report and the alleged victim was either charged or told that she (there were no male victims amongst these 17 cases) would be charged unless she dropped the complaint. While this represents only a fraction of the sample, the findings will show a much larger proportion of cases where police were confident, or reasonably confident, that the allegations were false but there was no attempt to institute charges against the alleged victim.

Heenan & Murray (2006, 24) specifically said that there were another 77 cases where it appeared that police were sure or suspected that the allegations were false, but where the outcome was classified as "no further police action". Adding 77 cases lifts the overall prevalence rate to around 11.6% (94/812).

There may also have been further false allegations in the "no further police action" category[28] and among the other categories described in the database, including "complaint withdrawn" and "case still ongoing", which together accounted for over a third of the cases.[29] And there may have been some false cases among the 15% of cases where charges were laid. The study did not canvas this matter.

Together with the various data limitations that afflicted the study, these uncertainties mean that it is not possible to devise a robust estimate of false reports, or an upper bound estimate, from the study.

The 1979 Philadelphia study



In a now quite dated study initiated by the Philadelphia Center for Rape Concern, McCahill, Meyer & Fischman (1979) examined a sample of 1401 alleged rapes and attempted rapes reported to Philadelphia General Hospital between April 1973 and June 1975. The study covered females of all ages. They were medically evaluated and in most cases were interviewed by the police. They were also invited to participate in a series of home visits and interviews by social workers and, for those accepting, interviews with psychiatrists as part of their aftercare. Just over half of the complainants (790) participated in at least one in-the-home interview with a social worker.

While the study focussed on the treatment of, and effects on, victims post-rape, the study also enabled some scrutiny of police classifications, including those allegations deemed by police to be unfounded. The study showed that police would sometimes decide to unfound cases for reasons other than having evidence that a crime did not occur. These included scepticism about the victim for reasons including drug addiction or prostitution, perceived victim precipitation and prior sexual contact with the perpetrator.

Of the 709 cases for which both police and social worker reports were available, the police classified 15% as unfounded (and may have disbelieved many others[30]), but only 3% were judged to be false allegations by social workers who sought to apply the definition that "investigation shows that no offense occurred nor was attempted" (p. 120). However, based on social worker assessments of the alleged victims' credibility, a larger number — 12.7% — were equivocal or suspicious.[31]

While this might be taken to provide an upper bound estimate of the prevalence rate, the study's limited survey response rate raises questions about the robustness of its estimates. Almost half of the females who made the initial 1401 rape complaints did not participate in interviews with social workers and thus were not included in the prevalence rate statistics.[32] The authors acknowledged that "the representativeness of the sample of victims interviewed, when compared to all victims reporting to the Philadelphia General Hospital, is difficult to assess, because of the unavailability of detailed information on those cases that [attempts to obtain an interview] were unsuccessful or [victims] refused to participate." The authors speculated that some of these complainants may have avoided participating "to ward off further abuse by the medical and criminal justice systems". (p. 11) This seems a reasonable hypothesis. Yet it seems equally reasonable to speculate that some complainants who made false allegations would also seek to avoid the additional interviews — and at a potentially much higher rate than those who made false allegations and agreed to the additional scrutiny.[33]

This makes it hard to draw firm conclusions about the prevalence of false allegations among the 1401 rape complaints made to the Philadelphia General Hospital, or even to specify a reasonably firm upper bound estimate.

The 1977 Toronto study

In another early study, Clark & Lewis (1977) examined a sample of rape cases in Toronto to provide clearer data on rape and how allegations were handled by police and other institutions, and to raise consciousness about problems with the system.

The authors asked the Toronto Police Department for access to its case files on all rape allegations it received in 1970, as well as permission to interview the complainants. Files in the form of General Occurrence Reports were provided for all the cases, although the authors noted that the quality of information contained in these reports varied. Identifying information on complainants, offenders and witnesses was redacted, reflecting police concerns to protect confidentiality. The authors were thus unable to interview complainants.

The authors excluded cases of attempted rape and rape of females under 14 years of age. There were no male complainants in the sample because, at the time, the relevant statute defined rape as a male-on-female act. The law also excluded what would now be considered spousal rape. This left the authors with a sample of 116 rape cases, around 7% of which involved multiple offenders and/or victims.

The police classified the completed cases as either founded or unfounded. Founded cases were those where police believed there was sufficient reason to lay charges. Unfounded cases had been dropped.

The study authors introduced a third category, "unfounded/possibly founded", for those cases the police had classified as unfounded but for which there was no material evidence for disbelieving the complainant's allegation of rape. The authors assigned cases to this category if they deemed that any of the following conditions were met:

- the police reason (stated or implied) for unfounding was the alleged victim's unsuitability as a witness, due for example to personal circumstances or qualities such as mental health problems that might limit her credibility
- the police reason for unfounding was a lack of solid corroborative evidence that would be acceptable in a courtroom (noting that, at the time, such evidence was effectively essential to secure a rape conviction)
- the alleged victim had withdrawn from the process
- there was independent evidence indicating a rape had occurred but the police unfounded the case anyway
- the police appeared unwilling to investigate a plausible complaint due to personal or other prejudice.

Of the 116 cases, the police had classified 74 as unfounded. The authors assigned 62 of these to their "possibly founded" category, leaving just 12 cases as in which the complainant recanted or there was other evidence that a rape did not occur. This yields a false allegation prevalence rate of 10.3%. This estimate fed into the 2016 Ferguson & Malouff meta-analysis.

Ferguson & Malouff included the 62 "possibly founded" cases in their listing of cases deemed as suspicious or equivocal by study authors. Adding these cases would yield an upper bound prevalence rate estimate of 63.8% (ie 74/116). However, while it is likely that some of these cases were indeed false reports, we do not see this as a meaningful upper bound estimate.

There are good reasons to believe that many, and potentially the bulk, of these cases were not false reports. (Our earlier discussion of the 2005 British study sets out several reasons, other than the innocence of the accused, as to why victims may withdraw, why insufficient evidence may be found and why cases may not be prosecuted.) The study authors themselves said "we believe that a rape had most likely occurred in 104 of the 116 cases we studied." (p. 57).

While it is not possible from the information in the study to know how many cases in this category were false, for each tenth of the 62 cases that were false, the prevalence rate would rise by 5.3%. So, for example, if one-in-ten of these cases were false, the overall prevalence rate would be 15.7%. Were the rate one-in-five, the prevalence rate would be 21.0%. Again though, specifying a reasonably firm and meaningful upper bound estimate for the sample in the study is not possible.

Conclusions on the high-quality studies' prevalence estimates

While the high-quality studies are more rigorous than many other studies in the field, our analysis demonstrates that they have important limitations. These include the way they classify cases as false, which leaves out many false and potentially false reports. Incomplete or poor-quality data, poor interview response rates and mathematical shortcomings also weaken some of the studies.

The prevalence rate estimates in those studies are more properly seen as lower bound estimates. The actual prevalence rate could be noticeably higher than the weighted estimate of 5% calculated (but caveated) by Ferguson & Malouff (2016), or the top of the 2–10% range from Lisak et al. (2010). However, the high-quality studies provide insufficient information to pinpoint the actual prevalence rate, or to devise meaningful upper bound estimates.

This reflects the inherent difficulties in separating fact from fiction in many sexual assault cases. It is doubtful that any study, largely regardless of how well it was resourced and conducted, could tightly estimate the prevalence rate.

While of necessity we have used arbitrary assumptions to provide some sense of how high the actual prevalence rates for some of the studies' samples could be, they are no more arbitrary, but we contend less unrealistic, than the equivalent assumption in the studies. The assumption that underlines their estimates is that none of the cases outside those classified as false under official crime reporting classification rules are in fact false.

Discussion and implications for the sexual violence literature

Our findings suggest a need to revise the contemporary position that the prevalence rate can be and is known within tight bounds and is very low. Importantly, nothing we have found challenges the view that the majority of sexual assault reports are true: contrary to some historical beliefs around rape, there is no credible evidence that women routinely fabricate sexual assault claims or that a substantial share of reports are false. But our findings point to more uncertainty about the actual number of false reports than the contemporary sexual violence literature recognises.

Our analysis in turn challenges the approach adopted by sexual violence researchers for classifying reports as false. Even if the approach is fit for official policing purposes, it is unfit for the very different purposes to which estimates derived from empirical prevalence studies are put.

The main problem is that the studies do not make any allowance for those reports that are false but have not been or cannot be demonstrated to be false. No scholar writing on sexual violence matters would suggest that the only valid or true reports of sexual assault are those that police find sufficient evidence to classify as "founded" or "detected" (or which the courts find "guilty"), as they would recognise that a good number of other reports were also true, but the victim withdrew or for other reasons there was insufficient evidence to warrant an arrest (or conviction). Yet by adopting the IACP approach to *classifying* reports as false (as distinct from the IACP *definition* of a false report), the studies effectively presume that the only complaints that are false are those where there would be sufficient evidence for police to classify them as false — when of course, just as with true reports, there are likely to be false reports in other categories.

Additionally, several specific classification choices made in the studies mean that, for reasons other than lack of confirmation of the falsity of the report, they omit reports that adhere to the IACP definition of a false report. Again, that *definition* is that a false report is a report of sexual assault where "no crime was committed or attempted". The definition does not limit who can lodge a false report to only the alleged victim; nor does it say anything about the motives and intent of the person lodging the report. Thus, false reports lodged by family and friends of the alleged victim are still false reports; and false reports lodged without malice are still false reports. Including these types of false reports in prevalence estimates is necessary to provide an accurate sense of what share of reports made to police are reports of an actual crime and what share are not, and thus to what extent complaints in general are valid. Yet,

some of the studies exclude false reports lodged by family, friends or other third parties from their prevalence estimates, and most exclude "baseless" reports and/or some others where a report of sexual assault was lodged but it transpires that a crime was not attempted or committed.

In a recent survey of the prevalence literature, Orchowski et al. (2020, 3) sought to explain and give credence to the studies' approach to classifying reports as false, including their use of the IACP classification rule, in the following way (with emphasis as in the original):

> Investigation into sexual violence may label accusations as "insufficient evidence", "unsubstantiated", or "baseless" if there is not enough evidence to conclusively prove that an assault was attempted or completed. A report of this nature does not meet the elements for classification as a crime but *still can be presumed to be a truthful report*. … [These claims] must therefore be distinguished from "false reports" of sexual victimization; such that the former are presumed to be a truthful accusation that has not been disproven, whereas the latter is proven to be an intentionally fabricated experience.

However, there is a logical leap in moving from the observation that an unproven accusation *can* be presumed to be truthful to implying that they *must all* be presumed truthful, with none false. Of course, as we have recognised, there is little doubt that the majority of accusations made to police are reports of actual sexual assault crimes, and it may well be desirable for police investigators to commence any individual case with the mindset that the accusation is *likely* to be true and to interact with complainants *as if* they believe the complaint to be true.[34] It is also fair to distinguish reports that have been confirmed false from those whose veracity or falsity has yet to be determined. However, it is an extreme and unjustifiable assumption to presume that no unproven accusation is false. And it is potentially misleading to promulgate prevalence estimates based on that assumption, as the high-quality studies effectively do, at least without properly communicating that assumption and its ramifications.

Some of the high-quality studies' authors, and other academics drawing on them, have not communicated these matters clearly. They have described the prevalence estimates without recognising that because they cover only confirmed false reports they omit others. Of those studies published this century, only Heenan & Murray (2006) and Spohn et al. (2014) acknowledged this point, and that there may thus be more false reports than their estimates captured.[35] Further, none of the studies question or explain the fitness-for-purpose of using the official crime reporting or case counting rules for their studies. Rather, the rules are simply asserted to be the correct way to classify cases as false or not, implicitly with an appeal to authority (of, for example, the British Home Office, the US Federal Bureau of Investigation and/or the IACP). They are then used in the studies, generally without relevant caveats.

Outside of the studies, academic discussion also often proceeds without justifying and/or caveating what is and is not covered by "false reports" in the studies, with academics themselves sometimes misunderstanding the estimates. For example, Weiser (2017, 53) assumed that the prevalence estimates of 10% or less from the high-quality studies meant that "at least 90% of reported sexual assault cases are truthful". In effect, the author wrongly assumed that the estimates captured all false allegations, and that any allegations that were not classified as "false" were "truthful".

This in turn implies a need to better explain, caveat or correct the terms used in prevalence studies. This could involve:

- relabelling prevalence estimates as "lower bound estimates" or as "estimates of *confirmed* false reports", as Ferguson & Malouff (2016) termed them
- designating new case categories or overarching groupings in prevalence studies, such as "possibly false" or "veracity not determined", to accompany the "confirmed false" category
- explicitly noting if and where baseless, third party, mistaken or other false reports are excluded from the prevalence estimates, and explaining the reasons for and ramifications of this.

On this latter point, scholars in the sexual violence field sometimes use prevalence estimates to try to answer more nuanced questions than just "what share of reports is false?". Such questions include whether people (mainly women) who present as sexual assault victims should be believed rather than mistrusted, and whether men in general are much affected by false allegations. It might be argued that these purposes warrant slightly different "slices" of the data. Box 4 discusses these issues and explores some options for alternative slices. In our view, however, given the significant uncertainty about the actual extent and nature of false reports, attempting to additionally manipulate the available data in this way may not shed much light on these questions.

Should the data be cut differently?

Leaving aside issues of measurability, prevalence rate estimates that delineate true and false complaints under the IACP definition could be used to draw inferences about the share of sexual assault reports that are valid and thus about the general credibility that criminal justice system personnel and the public should attach to complaints.

However, there are two more-nuanced questions, which frequently arise in the sexual violence literature, that could in principle warrant different cuts to prevalence estimate data.

First, many scholars who calculate or use prevalence estimates aim to answer the question: in general, how much credibility should criminal justice system personnel and the public attach to the people who report that they have been sexually assaulted? This question arises particularly in relation to females, who have historically been mistrusted in relation to rape allegations due to patriarchal norms by those (mainly men) in the criminal justice system (Jordan 2004, 30-33; Weiser 2017, 47).

To focus on this issue, one approach might be to remove third party reports, such as those filed by family and friends, from the data. It might be argued that unintentional or mistaken reports should also be removed as such reports are "honest mistakes" rather than warranting mistrust in the complainant. This would leave the estimates as "fabricated" (rather than "false") report prevalence estimates. The high-quality prevalence studies often take these approaches.

There are counterarguments. Briefly, in relation to unintentionally false reports, the propensity to make even honest mistakes when reporting is still relevant to determining the level of trust to place in what complainants say. In relation to reports by third parties, we understand that these will often have been triggered by the alleged victim making the allegation to the third party, or not correcting an assumption made by the third party, and then not interceding when that person decides to file that allegation with police (as happened in, for example, many of the "(later retracted) false allegations" described by Jordan (2004, 50)). In these cases, although not making the report to police, the alleged victim could still be considered responsible for its falsity.

The larger issue is that, even if these matters could be resolved, they would likely have a marginal impact on the estimates relative to the base number of false reports, which we cannot estimate with any precision.

Second, a concern in the sexual violence literature is that reform efforts to improve victims' access to justice can be derailed by "male-centred overconcern" that false allegations too often harm alleged perpetrators (Weiser 2017, 54). In the context of a debate in the British Parliament in 2010, Kelly (2010, 1349) said that "[t]he spectre invoked by media and politicians of large numbers of men being falsely accused and suffering the ignominy of public exposure is not borne out by these data". This message is premised on the view that cases in which a named individual is falsely accused are rare. To this it is added that those cases in which a falsely accused is arrested or charged are rarer again. Weiser (2017, 53-54) summarises the empirical evidence for this view as follows:

> Interestingly, there is a small amount of research that suggests that false reports are more likely to follow a stereotypical rape narrative that a stranger attacked and sexually assaulted the victim, making identification of a suspect unlikely (Lonsway 2010; Spohn et al. 2014). Spohn et al. (2014) suggested that false reporters seem to think that their account will be more believable if it conforms to the stereotypical stranger assault. Additionally, the supposed victim named a suspect in only 18% of the cases that Kelly (2010) identified as false reports, and arrests were made in only six cases (2.8%) that were later deemed to be false reports, and the victim was charged in only two of those cases (ie 0.9% of cases that police later classified as false). Other research also suggests a similar pattern in which false report cases are more likely to include a stranger as the perpetrator and no suspect was ever named by the alleged victim (Heenan & Murray 2006; Lonsway 2010).

To focus on this issue, in principle prevalence estimates could be sliced to remove those reports without named offenders. The resultant "prevalence rates of false reports with named offenders" estimate would then speak most directly to the issue raised by Kelly.

This might be of limited utility in practice, however, because we do not know and cannot estimate with any confidence how many false allegations of this type there are. The view in the literature — that many of those who make false reports do not name an offender — is based largely on assessments of those cases determined to be false in accordance with the IACP classification rule and other classification choices made in the high-quality studies. That view is not informed by the potentially many more false allegations that are not classified as such. For example, while Kelly found that only a limited number of the cases she classified as false involved named suspects or led to arrest or charge, she did not consider whether the potentially much larger number of false cases that her study did not capture might more frequently involve a named suspect, who then might also have been investigated, arrested or charged. As noted earlier, there could easily be up to around 5 times as many false cases in the 2005 study data than Kelly and her colleagues estimated. We simply do not and cannot know what share of these cases involved an identified suspect, arrest and charge.

References

Archambault, J. & Lonsway, K. 2021. 'Clearance methods for sexual assault cases', End Violence Against Women International, (first published July 2007; updated July 2021), https://evawintl.org/wp-content/uploads/Module-13_Clearance-Methods-Module.pdf (accessed 14 November 2021).

Clark, L. & Lewis, D. 1977, *The price of coercive sexuality*, Toronto Canada, Women's Press.

De Zutter, A., Horselenberg, R. & van Koppen, P. 2017, 'The prevalence of false allegations of rape in the United States from 2006-2010', *Journal of Forensic Psychology*, 2, 119-126.

Ferguson, C. & Malouff, J. 2016, 'Assessing police classifications of sexual assault reports: A meta-analysis of false reporting rates.' *Archives of Sexual Behaviour*, 45(5), 1185–1193.

Heenan, M. & Murray, S. 2006, 'Study of Reported Rapes in Victoria 2000-2003', Summary Research Report, Office of Women's Policy, Government of Victoria.

Henriques, R. 2016, 'An Independent Review of The Metropolitan Police Service's Handling of Non-recent Sexual Offence Investigations against Persons of Public Prominence', https://www.met.police.uk/SysSiteAssets/foi-media/metropolitan-police/other_information/corporate/mps-publication-chapters-1---3-sir-richard-henriques-report.pdf (accessed 14 January 2022).

IACP (International Association of Chiefs of Police) 2005, 'Investigating Sexual Assaults, Concepts and Issues Paper', National Law Enforcement Policy Center, Originally published November 1999, revised July 2005.

Jordan, J. 2004, 'Beyond belief? Police, rape and women's credibility', *Criminal Justice*, 4(1), 29-59.

Kelly, L., Lovett, J. & Regan, L. 2005, 'A gap or a chasm? Attrition in reported rape cases', Home Office Research Study, London, England.

Kelly, L. 2010, 'The (in)credible words of women: False allegations in European rape research', *Violence Against Women*, 16(12), 1345-1355.

Lisak, D., Gardinier, L., Nicksa, S. & Cote, A. 2010, 'False allegations of sexual assault: an analysis of ten years of reported cases', *Violence Against Women*, 16(12), 1318-1334.

Lonsway, K., Archambault, J. & Lisak, D. 2009, 'False Reports: Moving beyond the issue to successfully investigate and prosecute non-stranger sexual assault', The Voice 3(1), https://evawintl.org/wp-content/uploads/FalseReports-Movingbeyondtheissue.pdf (accessed 2 June 2022).

Lonsway, K. 2010, 'Trying to move the elephant in the living room: Responding to the challenge of false rape reports', *Violence Against Women*, 16(12), 1356-1371.

McCahill, T., Meyer, L. & Fischmann, A. 1979, *The aftermath of rape*, Massachusetts, Lexington Books.

Minter, K., Carlisle, E. & Coumarelos, C. 2021, '"Chuck her on the lie detector": Investigating Australians' mistrust in women's reports of sexual assault', Research Report, November, ANROWS, https://20ian81kynqg38bl3l3eh8bf-wpengine.netdna-ssl.com/wp-content/uploads/2021/10/ANROWS-NCAS-Mistrust-of-reporting-SV.pdf (accessed 2 June 2022).

Murray, S. & Heenan, M. 2012, 'Reported Rapes in Victoria: Police Responses to Victims with a Psychiatric Disability or Mental Health Issue', *Current Issues in Criminal Justice*, 23(3), 353-368.

National Sexual Violence Resources Centre undated, https://www.nsvrc.org/statistics (accessed 31 May 2022).

Orchowski, L., Bogen, K., & Berkowitz, A. 2020, 'False reporting of sexual victimization: prevalence, definitions, and public perceptions', in Geffner, R., White, J., Hamberger, L., Rosenbaum, A., Vaughan-Eden, V. & Vieth, V. (eds.), *Handbook of Interpersonal Violence Across the Lifespan*, Springer, Cham.

Rumney, P. 2006, 'False allegations of rape', *Cambridge Law Journal*, 65(1), 128-158.

Saunders, C. 2012. 'The truth, the half truth and nothing like the truth: Reconceptualizing false allegations of rape', *British Journal of Criminology*, 52(6), 1152–1171.

Spohn, C., White, C. & Tellis, K. 2014. 'Unfounding sexual assault: examining the decision to unfound and identifying false reports', *Law & Society Review*, 48, 161-192.

Weiser, D. 2017, 'Confronting myths about sexual assault: A feminist analysis of the false report literature'. *Family Relations*, 66(1), 46-60.

Notes

1 The authors have had more than three decades experience each as researchers and policy analysts with Australian Government agencies. Our work has covered an array of social, cultural, economic and legal issues, and we have extensive experience in handling statistics and critiquing empirical research. We now work as independent researchers. In compiling this paper, we sought clarifications and feedback from a number of sexual violence researchers, including authors of all the recent (post-2000) high-quality prevalence studies, at various points during the paper's gestation. We also circulated a full draft to them in December 2022, with a further invitation for comments, corrections or suggestions. We thank those academics who responded, and welcome further feedback.

2 Like several papers in the literature, we generally use the terms "sexual assault" and "rape" interchangeably. Legal definitions of the terms differ slightly across jurisdictions.

3 Prominent researchers associated with the development of this consensus include Prof Liz Kelly, Dr Kim Lonsway and Dr David Lisak, all of whom delivered papers to a Symposium on False Allegations of Rape, published in the journal *Violence Against Women* in 2010 (Kelly 2010, Lonsway 2010, Lisak et al. 2010). In her review of the research, Lonsway (2010, 1358) argued that "Because the studies highlighted in [Dr Lisak's] review meet generally accepted standards for social scientific research, they instil confidence in the credibility of their estimates — which range from 2% to 8% or 2% to 10%, depending on the specific studies included. In fact, the divergence of methods means that the convergence of findings is especially noteworthy. At this point, there is simply no way to claim that 'the statistics are all over the map'. The statistics are now in a very small corner of the map." Lonsway (2010, 1366) indicated that the true number is likely to be closer to the bottom of the range, a view said to be shared by Kelly (referenced in Minter et al. 2021, 13). More recent commentary by academics in the sexual violence space affirms the consensus view that false reports are rare (see, for example, Weiser 2017, 51-53; Orchowski et al. 2021, 4-6).

4 As well as including the 1992 British study by Grace, Lloyd & Smith and the 2005 British study by Kelly, Lovett & Regan in their selection of credible studies, Lisak and his colleagues also summarised a 1999 British study by Harris & Grace. The 1992 study estimated the prevalence rate at 8.3%, and the 2005 study estimated the rate at 2.5–3%, while the prevalence rate estimate in the 1999 study, which was based purely on (unscrutinised) police classifications, was 10.9%. Lisak et al. (2010, 1330) appear to have discarded the latter estimate when arriving at their range of credible estimates of 2–10%. Ferguson & Malouff included neither the 1992 nor the 1999 British studies in their meta-analysis, but retained the 2005 British study by Kelly, Lovett & Regan, which will have reduced their estimated prevalence rate.

5 In their review of the literature, Orchowski et al. (2020) identified only one new prevalence study since Spohn et al. (2014), namely a review of the FBI Uniform Crime Reporting program by De Zutter et al. (2017). Among other things, this review found that the number of rapes classified as false or baseless by US police forces has fallen since the adoption of tighter classification guidelines. However, the study did not involve independent assessments of cases classified by police or consider the scope for there to be false cases in other categories.

6 Under the heading "What is the actual definition of a false report?", Lonsway, Archambault and Lisak (2009, 4) state: "Although many people have different ideas about exactly what constitutes a false report, the most reasonable definition is that: A false report is a report of a sexual assault that did not happen (ie, it was not completed or attempted)." This is the definition embodied in the IACP statement. These authors then differentiate between the *definition* and the *classification rule* used by police and prosecutors (and in the high-quality studies) by saying "While we might all agree with this simplistic definition of a false report, people have different ideas about exactly when they can decide that the sexual assault did not actually happen. … In reality, investigators and prosecutors cannot determine that the sexual assault did not happen, simply because they suspect that the report is false, view it with suspicion, or because the victim changes his or her account of what happened …. [or] because the victim lacks credibility. … Rather, investigators and prosecutors must base all final judgments on the findings from a thorough, evidence-based investigation. The determination that a report is false can then only be made when there is sufficient evidence to establish that the sexual assault did not happen (was not completed or attempted)." Note that while Lonsway, Archambault and Lisak (2009, 4-5) explicitly separate the definition of a false report from the rules for classifying reports as false, the other high-quality studies, while adhering to the IACP policy for defining and classifying false reports, generally do not raise or dwell on this matter.

7 A sexual assault report's "falsity" or "truthfulness" can be interpreted in different ways. Saunders (2012) differentiates between allegations of sexual assault, accounts of the alleged events, and statements about particular details within those accounts. Saunders notes that any particular statement may be factually true (or not); an account can be taken to be truthful (or at least generally true) if it reasonably depicts an event that occurred, even if some details within the account were not true; and an allegation of sexual assault could be labelled as true even if the account of the circumstances and manner in which the crime occurred (including the identify of the assailant) were not truthful. Saunders found that police in her study might describe reports containing untrue statements and accounts as false, even though the allegation was of an actual crime; whereas sexual violence academics would be inclined to label such a report as truthful or, at least, not false. In this paper, we use the terms true and false largely in line with the academic interpretation. That said, we recognise that allegations that accuse someone other than the real perpetrator of an actual sexual assault would be classified by many people as false.

8 Ferguson & Malouff (2016, 1188) noted that, beyond the way they classify allegations as false, the high-quality studies have a range of limitations. Problems included missing, incomplete or low quality police case files and other data problems; limited responses to attempts to interview complainants; for three of the seven studies, a lack of information on how many cases were in doubt in addition to those confirmed to be false; and sampling differences (for example, some did not include male complainants). The ways the individual cases are evaluated in the studies may also leave the door open in some cases for evaluators' subconscious biases to influence results. We also note that some of the studies are now quite dated, and mathematical shortcomings affect some estimates.

9 Kelly et al. (2005, 50). The study's prevalence rate estimate is sometimes said to be 2.5% and sometimes said to be 3%. (Note that 67/2643 = 0.0253 and 67/2284 = 0.0293.) Kelly et al. (2005) use the 3% figure in the text, and Sphon et al. (2014, 165) and Weiser (2017, 52) also used this figure. However, several other academics have used the lower 2.5% estimate (see, for example, Ferguson & Malouff 2016, 1189; Lisak et al. 2010, 1326; Lonsway et al. 2009, 3; and Orchowski et al. 2020, 5). To set the numerator and denominator on a like basis, we consider that the higher estimate is the more appropriate (although, for various other reasons explained in the text, we consider that both estimates are significant underestimates of the actual prevalence rate.)

10 We wrote to the authors in October 2021, and again in both May and July 2022, inviting clarifications, confirmations and/or corrections on the matters in this section. We also circulated a full draft of the research paper for comment to them in December 2022. In response to the first approach, Prof Kelly replied that the authors could not respond at that time due to heavy work demands but would reply at a later date. We have not received any further response. Thus, while our analysis has been vetted by several colleagues, we have not had the benefit of feedback from the authors to either confirm or correct our understanding and analysis of their study.

11 As a share of all complaints reported to police, the prevalence estimate would be (67x216/144)/2643 = 0.0380. However, given that the numerator (67) is drawn from the subset of all complaints for which outcomes are known, the denominator should be the 2284 complaints for which outcomes are known for the comparison to be on a like-for-like basis.

12 Under the second pathway described in box 2, the appropriate calculation would be (77x216/144)/2284 = 0.0506; whereas under the third pathway, the calculation would be (77x216/154)/2284 = 0.0473.

13 Kelly et al. (2005, 46) noted that, in 12 cases in the "no evidence of assault" category, police made an initial coding of "suspected rape" when the presumed victim was too drunk or distressed to communicate what had happened, but evidence such as injuries or torn clothing suggested that a rape may have occurred. However, when the presumed victim was again able to communicate, it became clear this this is not what had occurred. These cases are not false reports to police, but nor should they be included in the denominator of false report prevalence estimates.

14 Calculated as (11x83/56 + (77/3)x216/144)/2284 = 0.0240. Note that including one third of the uncertain cases implies that, together with the earlier inclusion of all probable and possible cases in the study authors' estimates, around 65% of the cases classified as false by police are being counted toward the false case total. This figure could be reached in a number of ways, including allowing for the chance that some of the cases labelled "probable" or "possible" by the study authors were not in fact false reports.

15 Calculated as (662x0.1)/2284 = 0.0290.

16 Calculated as (633x0.1)/2284 = 0.0277. If the 40 victim withdrawals during later phases were included, the relevant calculation would be (673x0.1)/2284 = 0.0295.

17 There are many combinations of category rate assumptions that would generate a prevalence rate estimate approaching 15%. As one example, if to all 11 non-complainant "no evidence of assault" reports were added half of the "uncertain" false reports and 10% of reports in the "insufficient evidence etc" categories, the "victim withdrawal" categories, the "acquittals" category and the "reasons unknown" category, the resultant additional prevalence rate estimate would be 10% (calculated as (11x83/56 + (77/2)x216/144 + 0.1x(662+673+104+136))/2284 = 0.1014). To this would need be added the adjusted lower bound estimate of 4.4–5.1%. (An overall prevalence estimate of 20% could be reached if, for example, the assumed share of false reports in the four latter categories were increased by a further 7–8 percentage points).

18 Spohn, White & Tellis (2014, 190) acknowledged that while they were provided with a (redacted) copy of each case file, they could "not know with any degree of certainty whether the information recorded in the case file was an accurate and unbiased report of what happened and what complainants, witnesses, and suspects said about the alleged incident". They accepted that this, together with the fact that they examined only cases classed as unfounded, "suggests our estimate of the prevalence of false allegations of rape may underestimate the actual rate".

19 Such a case was classified as false "only if there was independent evidence that a crime did not occur and there was no evidence in the case file that the complainant's recantation was motivated by fear, pressure or a belief that prosecution would not be in the complainant's best interest" (Spohn, White & Tellis 2014, 172).

20 Baseless reports are allegations that are initially recorded as crime reports but where, on investigation, it becomes apparent that the legal elements of the criminal offense were not met. They may be truthful accounts, and involve a named perpetrator, but the actions reported were not illegal (or, at least, did not meet the elements of the particular crime reported). For example, a women might report a sexual act that was unwanted when the act did not meet the requisite criminal threshold for force, threat of fear to be classed as a sexual assault (Archambault & Lonsway 2021, 20).

While crime reporting rules require police to differentiate between false and baseless reports, a question is whether such reports would be considered "false reports", either in the common sense of the term or against the IACP definition that "no crime was committed or attempted." If so, they could be added to the numerator in prevalence estimates. An alternative, albeit not clearly a better one, may be to remove them from the denominator on the premise that they were mislabelled as reports of a sexual assault crime.

21 Lisak et al. (2010, 1328) explain that this classification was applied "if a report lacked basic information (e.g., neither the victim nor the perpetrator was identified, and there was insufficient information to assign a category)". As noted, there is a case to remove these 19 cases from the denominator when calculating the prevalence rate. Alternatively, it could be appropriate to inflate the numerator in proportion to the rate for the other categories combined, on the assumption that the rate of false allegations in this category is around the same as the rate across the categories for which assignments were possible.

22 We approached the authors for a response to our analysis but Prof Gardinier replied that she no longer had access to the study files. We have not been able to contact and/or have not received replies from Dr Lisak or the other authors.

23 The percentage figures are the share of the 1984 police cases attributed to each category. The 1984 case figure was derived by removing unspecified cases from the starting sample of 2059 police cases. Source: page 8 of the MAD Data Findings slide document supplied by Dr Lonsway.

24 The EVAWI website (https://evawintl.org/lessons/decide-on-informational-vs-crime-report/, accessed 13 November 2022) provides three examples of the use of informational reports: where a complainant reports a situation that does not meet the criteria for sexual assault; where law enforcement records information about a suspicious individual or incident (which does not at that stage meet the elements of a crime); and where a third party reports a sexual assault allegation but the alleged victim denies the allegation and/or is unable or unwilling to provide the information necessary for police to complete a crime report. While the first category is not technically reports of sexual assault, an informational report of the third type (and possibly also of the second type) could be classed as a report of sexual assault, albeit one whose veracity is uncertain given the absence of corroboration. It could be argued that these two types of informational reports should be retained in the denominator for false report prevalence estimates, although we have no basis for knowing what share of informational reports are of this type. Moreover, if these reports were to be counted in the denominator, it would seem appropriate to also adjust the numerator to reflect that some share of these allegations might well be false.

25 Calculated as 140/(1984x(1-0.179)) = 0.0859. With the additional removal of baseless reports from the denominator, the calculation becomes 140/(1984x(1-(0.179+0.085)) = 0.0959.

26 The main publicly-available document on the study appears to be Lonsway, K., Archambault, J. & Lisak, D. (2009), which is an article in "The Voice" that covers several matters (such as advice for police and prosecutors) in addition to describing and explaining the results of the study. (A version is also published on the USNSVRC website.) EVAWI's website also contains several short documents about the study, including on 'Research Methods' and 'Basic Data Findings' — the latter containing powerpoint slides. However, the authors were unable to obtain a 2008 unpublished manuscript that analysed the data from the MAD project (as referenced in Orchowski et al. 2020).

27 As noted earlier, Spohn et al. (2014, 186) reported that a number of Los Angeles detectives indicated that they would sometimes clear cases by exceptional means or keep cases open, even if they believed the report was false or baseless. The misclassification or inconsistent classification of cases by police is recognised as a common problem. However, we speculate that with the emphasis given to rigorous case classification as part of the training provided to participants in the EVAWI study, the police agencies involved in this study may have been less likely to clear false cases by exceptional means or with other inappropriate categorisations.

28 After stating that there were 77 cases in the "No further police action" (NFPA) category that the police were confident or reasonably confident were false, Heenan & Murray (2006, 24) added that in 44% of cases that resulted in NFPA, the police "recorded information about the case in neutral terms without expressing a view about the veracity of the allegations". This may mean that a share of these additional NFPA cases could also have been false.

29 The Heenan & Murray (2006) document is designated as a "Summary Research Report". We have found it difficult to divine from the summary report whether the 77 suspicious/equivocal cases were drawn solely from the category "No further police action", or whether they were drawn as well from other categories. We have been unable to reach the study authors to obtain clarity on this and some other matters.

30 McCahill, Meyer & Fischman (1979, 110) noted that the police could effectively unfound cases by applying a nonoffense charge, stating "Adding cases given a nonoffense charge to cases that were actually marked as unfounded in the current study indicates that the Philadelphia police disbelieved and turned away over 27.2% of the women who alleged rape."

31 Social workers were asked to evaluate the credibility of each alleged victim's report of the incident. Of the 577 cases evaluated, they rated 73 as "partially incredible". The authors emphasised that this classification indicates only that the alleged victim's story contains several questionable elements, which are properly reserved for courtroom scrutiny and are not conclusive proof that no rape actually occurred (McCahill, Meyer & Fischman 1979, 113 & 116).

32 McCahill, Meyer & Fischman (1979, 12) checked the demographic characteristics of complainants who agreed to be interviewed by social workers against those who did not and found few significant differences between the groups, apart from colour (with whites less likely to participate than blacks), although even this difference was not considered sufficiently strong to prevent generalising from the research group of 790 to the total sample of 1401. The demographic characteristics compared would not shed light on the propensity of the different groups (interviewees vs non-interviewees) to make false allegations.

33 A further complication is that, of the 790 complainants interviewed by social workers, police files could not be found for more than 80 of the cases (thereby reducing the sample for the prevalence estimates to 709). The authors acknowledged that factors that made police more likely to "lose" files included where the victim later acknowledged to a social worker that her rape allegation was false and where social workers believed that elements of the victim's story were not credible. It is thus possible that police simply "lost" some cases they believed were false reports (McCahill, Meyer & Fischman 1979, 112-113).

34 There are some differing views on the presumptions police should adopt regarding the veracity of complaints, and the language they should use to describe complainants. See, for example, Henriques (2016, 13-28). In this paper, we have at times used the term "victim" rather than "alleged victim" or "complainant".

35 Although Kelly and her colleagues, in discussing categories such as insufficient evidence, victim withdrawal and acquittals, did not mention the possibility that some of the cases in those categories were false reports, the authors stated in the final section of their study that "There are false allegations, and possibly slightly more than some researchers and support agencies have suggested. However, at maximum they constitute nine per cent and probably closer to three per cent of all reported cases. An overestimation of the scale by police officers and prosecutors feeds into a culture of scepticism…" (Kelly et a. 2005, 83). This might suggest the Kelly and her colleagues had some clue that there could be more false reports than their study estimated. While unclear, the 9% figure may be a nod to some police-based estimates (although on page 47 the study uses a figure of 8% for false reports designated by police as a share of all reports to police). However, the implication of the statement is that the study's 3% estimate, although perhaps not perfectly precise, is far more accurate.