

Have Chinese Researchers Found A Link Between Abortion And Breast Cancer?

A zombie has risen from the grave and is shambling its way around the websites of the anti-abortion lobby and stirring up the natives, so it's time for me to dig out my trusty [metaphorical] shotgun, track it down and blow its head off.

Yes, the Abortion-Breast Cancer (ABC) Hypothesis, which proposes the existence of a link between induced abortion and an increased long-term risk of breast cancer - in women, obviously - has resurfaced thanks to a Chinese meta-analysis of 36 cohort and case control studies published a couple of weeks ago in the journal "Cancer Causes & Control" which purports to show a statistically significant increase in breast cancer risk associated with induced abortion coupled with a dose-response effect, i.e. the risk increases with the number of induced abortions.

Unsurprisingly, this has got the anti-abortion lobby frothing at the mouth because they mistakenly believe it puts one of their favourite arguments for trying to scare the living shit of women who might be contemplating a termination back in play as indicated [here](#) by Life's press statement on the subject:

Major study in China shows significant risk for breast cancer in women who have abortions

The national pro-life charity LIFE has called for the abortion breast cancer link to be taken seriously after a major new study from China revealed significant risks for women who have induced abortions.

In what is probably the most powerful evidence brought to support the suggested link between induced abortion and breast cancer, the study published in the prestigious medical journal Cancer Causes Control, concluded that just one abortion increases the risk of breast cancer by 44 percent. For women who have had two abortions the risk rises to 76 percent and then almost doubles after three or more abortions.

But is this new study - Huang et al (2013) - really all that Life and other anti-abortion groups think it's cracked up to be?

Although I wrote a series of five lengthy articles last year, looking in detail at the ABC hypothesis, its provenance and the evidence for and against it, rather than refer people back to those articles before moving straight on to look at this this new Chinese I think it better to give everyone a more easily digestible recap of the main issues. So, to begin with, I'll explain what the ABC hypothesis actually is and then review the two key studies in the field prior to this new paper – Brind et al. 1996 and Beral et al. 2004 – so that's everyone's up to speed when we come to assess whether or not this new study genuinely does put the ABC hypothesis back on the table.

What is the ABC Hypothesis?

To begin with, there is good evidence to indicate a link between women's long-term risk of developing breast cancer and their reproductive history, and in particular their cumulative exposure over to time to ovarian hormones, particularly oestrogen. Oestrogen, which is produced by the ovaries, acts to stimulate cell growth in breast tissue and if, over the course of a woman's lifetime that cell growth gets out of control then, unfortunately, you have yourself a case of breast cancer.

A woman's reproductive history is one of many different factors that can influence their risk of developing breast cancer later in life. Factors such as the early onset of puberty or a late menopause can increase this risk by increasing women's overall exposure to these hormones, while pregnancy, childbirth and breast feeding can lower the risk by reducing the total number of reproductive cycles a woman undergoes over her fertile lifetime - and if you're at all curious and want to understand more about this relationship then, as ever, [Cancer Research UK has an excellent guide](#) to the various risk factors associated with breast cancer that is well worth reading.

So, what does all have to do with abortion?

As already noted, pregnancy, childbirth and breastfeeding are amongst the factors that can influence reduce women's long term risk of developing breast cancer, as indeed is the age at which women start having children and the number of children they carry to term. According to estimates based on breast cancer incidence rates during the 1990s, the cumulative incidence of breast cancer in Western societies would be more than halved if patterns of childrearing and breastfeeding were the same those in developing countries at that time, where women were having an average of 6.5 children and would typically breastfeed their offspring for around two years, rather than six to eight months, which about average for the US, UK and other developed countries.

Abortion is, of course, just one of a number of ways in which women in developed societies are able to control their fertility, choose when to, and when not to, have children and limit their family size. It is not an ideal way of doing these things - contraception is of course a much better option – but it is a necessary and important part of the overall package of reproductive rights that give women control over their own fertility and, ultimately, their own bodies.

So, by delaying the age at which they begin child rearing and limiting their family size women are, strictly speaking, increasing their long term risk of developing breast cancer, but of course abortion is not the only way this is done and, in any case, the choice women make about when to have children, etc. are based on many more considerations than just the possibility that their choice might influence their risk of developing breast cancer thirty years or so down the line.

So if that were all there was to the ABC hypothesis then it wouldn't be an issue that attracted too much attention, so why the additional focus on induced abortion?

Pregnancy throws up an additional layer of complexity to the overall picture. Having noted that pregnancy can act to reduce women's cumulative exposure to ovarian hormones by reducing the overall number of reproductive cycles they experience over their fertile lifetime, it also has the opposite effect, particularly during the early stages of a pregnancy, when women get an extra hit of oestrogen which their body needs to help sustain the pregnancy and support placental development. The upshot of this is that pregnancy, itself, causes a short-term increase in women's risk of developing breast cancer, one which has been estimated to last for around ten years after they became pregnant – but before you let that worry you, for most women that all happens at an age when their overall risk of breast cancer is so low that the transient increase in risk associated with pregnancy really doesn't make a lot of difference. It's only if a woman's risk of developing breast cancer is significantly higher than normal for genetic reasons that the increased short-term risk associated with pregnancy is of any potential significance.

Against this, some researchers have also argued that pregnancy also has the effect of inducing genetic changes in breast tissue in women, changes which confer an additional degree of protection against breast cancer. The theory here is that the genetic changes in breast tissue that occur in the latter stages of pregnancy to enable women to produce milk also, over the long term, make that

tissue more resistant to cancer, and as this process of cell maturation only occurs after a woman has been pregnant for 32 weeks, the interruption of a pregnancy before this point by way of an induced abortion leaves women all the pitfalls of the extra hit of oestrogen they get early in pregnancy but none of the protective benefits of the changes in breast tissue which are accrued only by carrying the pregnancy to term.

It is also argued that this effect is most marked where a woman carries her very first pregnancy to term but diminishes in size with each subsequent pregnancy. So, if a woman terminate her first pregnancy in order to delay motherhood, she not only loses the cancer-protective effect of carrying that pregnancy to term but that loss is pretty much a permanent one as any benefits accrued from having children further down the line will not make up for those lost when she had her termination.

On that basis it's argued that induced abortion serves as an independent risk factor for breast cancer in later life, over and above the general influence that patterns of childrearing have on that same risk. However, it has to be said that the biological evidence for this particular theory is currently rather short of being conclusive - when I looked in detail at the theoretical foundations of the ABC hypothesis, last year, neither of the two most recent papers I could find on PubMed that looked at the relationship between breast cancer risk and the pregnancy related changes in breast tissue [Baer et al. 2009 & Ramakhrisnan et al. 2004] offered any significant support for that theory.

Okay, so that's the ABC hypothesis and although it at least biologically plausible the evidence base behind it is, for the time being, best described as inconclusive bordering on non-existent, not that this prevents anti-abortion activists from promoting it as it were a proven fact.

Researching the ABC Hypothesis

Having explained what the ABC hypothesis is, it's now time for us to move on and start looking at the epidemiological research evidence from to see whether, and to what extent it actually supports that hypothesis, but before I start diving into the research literature I need to introduce a few basic research concepts which you'll need to get to grips with in order to make proper sense of the rest of this article.

In the absence of direct biological evidence to support the ABC hypothesis, the next best route towards obtaining evidence to test its validity, or otherwise, lies in epidemiological studies which look at the incidence of a specific disease or condition in a particular population and then try to identify what it is that makes the people who have this disease/condition different from the ones who don't in the hope that this will provide insights into its causes and/or factors that can influence people risk of contracting it.

The vast majority of ABC studies are based on one of three different research methodologies:-

In case-control studies, researchers take a relatively small group of women who have already been diagnosed with breast cancer (the case or study group) and compare them to a carefully matched group of women who haven't (the control group) the aim being to ensure that the two groups are, in most respects (e.g. age, ethnicity, number of children, etc.) as similar to each other as possible. The vast majority of case-control studies that seek to investigate the ABC hypothesis use a retrospective self-report design in which the women in both groups are asked to complete a questionnaire about their sexual and reproductive history and the data from these questionnaires is then analysed to identify any significant differences in the responses received from the two groups that might indicate differences between them that could help to explain why the women in the study group, but not the control group, have developed breast cancer.

In cohort studies, researchers start with a healthy population of women and follow them over a period of years, often anything from 10 years upwards, recording changes in everything from their study participant's lifestyles to their health outcomes and building up a large dataset that researchers can dip into and extract information from in order to investigate whether or not a relationship exist between different lifestyle factors, behaviours and/or life choices and any subsequent health outcomes. If you've ever watch any of the '7-Up' or 'Child Of Our Time' series of programmes these are essentially cohort studies albeit on a much smaller scale, in terms of the numbers of participants, than most epidemiological cohort studies.

Record linkage studies can follow either a case-control or cohort design and obtain their information on the medical histories of study participants from centralised clinical registries rather than self-report questionnaires. Studies of this type are, however, relatively rare as they can only be conducted both if centralised registries containing relevant data exist and if the data is held in an anonymise format that permits data from different registries to be accurately linked together.

Where case control studies typically use a retrospective design in which data on lifestyle and other factors of interest to the researchers is gathered after the people in the study group have been diagnosed with a particular condition (breast cancer in the case of ABC studies), cohort studies will often, but not always, use a prospective design in which this information is gathered before particular conditions or diseases emerge in the study population, making this type of study less prone to the problem of under reporting and recall/response bias, an issue we'll come to in more detail shortly. Record linkage studies are, of course, even less prone to these issues but not totally immune. A national abortion registry will, for example, only record abortions carried out legally in registered clinics in a particular country, but won't capture data where a woman has an illegal termination or travelled to another country to undergo such a procedure, so under reporting is possible but would typically be at a much smaller scale than occurs in self-report studies.

Although large scale prospective cohort studies tend to provide stronger, better quality evidence and are much less prone to confounding and bias, running them is much more expensive and time consuming, which is why the majority of ABC studies tend to be smaller case-control studies that lack the statistical power and reliability to provide definitive evidence for or against a link between abortion and breast cancer, and this bring a fourth type of study into play – meta-analysis.

Meta-analysis is simply a set of statistical methods and technique which enable researchers to combine a number of smaller, underpowered research studies into a single, much larger and more powerful study which, hopefully, will provide much better evidence than an individual study is capable of providing in isolation. At its best, meta-analysis is an extremely useful and powerful set of research methods but it's an approach that not without its problems. In particular, without putting in a lot careful work into assessing study quality and controlling for variations in study design and other potential sources of confounding and bias, much of which can only be done effectively if researchers have access to the original data used in studies and not just their published results, meta-analyses can easily fall prey to the 'garbage in, garbage out' problem in which pooling together data from a lot of small, poorly designed and unreliable studies leads to the publication of nothing more than a much larger, but still poorly design and unreliable study.

The three studies that I'm going to look at in more detail; Brind et al. 1996, Beral et al. 2004 and the new Chinese study, Huang et al. 2013 are all meta-analyses of data extracted from other published ABC studies, and I'll start with the papers by Brind and Beral as, together, these not only represent the state of play up to the point at which this new Chinese study found its way into print but also help to illustrate much of what needs to be looked in evaluating whether or not the paper by Huang et al. is capable of living up to the hype it's receiving from anti-abortion activists.

Brind et al. 1996.

Brind's meta-analysis was published in 1996 in the British Medical Journal's Journal of Epidemiology and draws together data from 28 papers published between 1957 and 1996, including papers translated from Japanese, Russian and Portuguese. In total, these papers cover 23 individual research studies of which 21 contributed data to Brind's analysis, supplying data on what appears to have been around 25,000 women.

Brind's headline finding from this study is a statistically significant odds ratio for breast cancer in women with a history of induced abortion of 1.3 compared to women with no history of induced abortion (Odds Ratio 1.3, 95% CI 1.2-1.4), which is roughly equivalent to an increase in relative risk of anything from 30-42% (it's impossible to be more precise due to the limited information presented in the paper) from which he goes on to assert that:

We are convinced that such a broad base of statistical agreement rules out any reasonable possibility that the association is the result of bias or any other confounding variable. Furthermore, this consistent statistical association is fully compatible with existing knowledge of human biology, oncology and reproductive endocrinology, and supported by a coherent (albeit incomplete) body of laboratory data as well as epidemiological data on other risk factors involving oestrogen excess, all of which together point to a plausible and likely mechanism by which the surging oestradiol of the first trimester of any normal pregnancy, if it is aborted, may add significantly to a woman's breast cancer risk.

This was not, however, a view shared by the majority of epidemiologists (Brind is a professor of biology and endocrinology at Baruch College, the senior college of the City University of New York) as noted here in [a 2003 article published by Discover magazine](#):

The vast majority of epidemiologists say Brind's conclusions are dead wrong. They say he conducted an unsound analysis based on incomplete data and drew conclusions that meshed with his own pro-life views. They say that epidemiology, the study of diseases in populations, is an inexact science that requires practitioners to look critically at their own work, searching for factors that might corrupt the results and drawing conclusions only when they see strong and consistent evidence. "Circumspection, unfortunately, is what you have to do to practice epidemiology," says Polly Newcomb, a researcher at the Fred Hutchinson Cancer Research Center in Seattle. "That's something Brind is incapable of doing. He has such a strong prior belief in the association [between abortion and cancer] that he just can't evaluate the data critically."

So what is that most epidemiologists have seen in Brind's paper that prompts them to regard his analysis as 'unsound'?

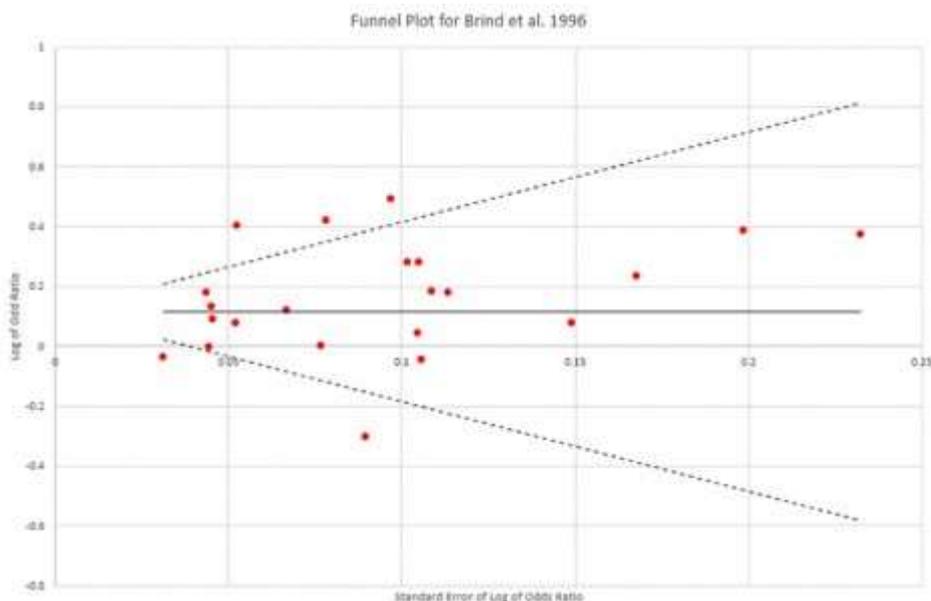
For one thing, Brind's paper lacks any kind of systematic review of the quality of the studies included in it. The paper does include a fairly lengthy narrative review, running to around a third of the total content of the paper, from which it's apparent that there are considerable variations in the design and quality of the studies included in Brind's analysis, which tends to suggest there may a significant risk of bias arising from inadequate and inconsistent controls for potential sources of confounding in individual papers, which Brind would be unable to correct in his own analysis. For example, several of the studies included in Brind's analysis do not distinguish between induced and spontaneous abortion (miscarriage) while at least one study makes this distinction only in relation to nulliparous

women (i.e. those who have never given birth) but not to women in the same data set who had given birth at some point in their lives. Allied to this, Brind's paper also lack any kind of statistical investigation for of study heterogeneity, so even if we suspect that either this may present a significant source of confounding/bias there is no way of evaluating the likelihood of this being a factor without going to the time of trouble of repeating Brind's analysis from first principles.

To put that observation into its proper context, at around the same time that Brind published his meta-analysis, Karin Michels and Walter Willett, of the Harvard School of Public Health published their own review of the available evidence from published epidemiological studies investigating the possibility of a relationship between abortion and breast cancer, one which covered substantially the same articles as were included in Brind's study. However, where Brind concluded from his analysis that he had found clear evidence of a link between abortion and breast cancer, Michels and Willett reached the conclusion that the studies published to that point in time were "inadequate to infer with confidence the relation between induced or spontaneous abortion and breast cancer risk" but that it appeared that "any such relation is likely to be small or nonexistent." [Michels & Willett, 1996]

The paper also lacks any kind of statistical investigation for [publication bias](#) although Brind does at least acknowledge this as a possibility, only then to dismiss it as a concern with the argument that because the suggestion of a possible link between abortion and breast cancer was somewhat controversial at the time it was most likely to be case that any publication bias would favour studies that reported no evidence of such a link.

However a funnel plot generated from the data reported in Brind's study tends to suggest that publication bias is not a significant issue although there a modest degree of asymmetry in the plot from a small number of lower quality studies reporting positive results finding their way into print that could introduce a slight positive bias but one too small to have a significant impact on Brind's results.



A more significant issue with Brind's paper arises from the fact that it appears that only two of studies included in his analysis used a prospective design, accounting for at best around 6.5% of the data included in his analysis, although a third study used record-linkage data and this appears to account for around another 12-13% of Brind's data, which mean of course that the vast bulk of the

data in this study, around 80%, comes from retrospective studies in which under reporting and recall/response bias may be a significant problem.

This is a significant issue in ABC studies, generally, and as such merits careful examination.

When researching potentially sensitive health issues, studies that rely on self-report measures, i.e. asking people to provide information via one of more questionnaires, will almost always result in a degree of under reporting on some questions. If we take abortion as an example, there will inevitably be some women who are extremely reluctant, if not wholly unwilling to disclose the fact that they have in the past had an abortion or, even if they do provide that information, disclose exactly how many abortions they've undergone if it's more than maybe one or two.

The range of factors that can influence the extent to which underreporting occurs in a particular study can be extremely wide, from general social attitudes and the degree to which a degree of social stigma is attached to certain things, like abortion, through to people's personal beliefs or the specific circumstances under which an abortion is carried out such as whether someone was underage at the time they fell pregnant, whether the pregnancy occurred as the result of rape or incest or, certainly in some older studies, whether or not the abortion was performed before the procedure has been legalised. As should be obvious, the extent to which any of these factors will influence some women to choose not to provide accurate information can vary considerably from country to country and even between different areas in the same country, as well as from generation to generation, between different ethnic and religious groups etc.

Underreporting of potentially sensitive issues does not necessarily create a significant problem for researchers.

In prospective studies, for example, data is gathered over a period time via questionnaires which ask study participants about a wide range of different health, some very sensitive and personal and some altogether more mundane and although some degree of underreporting is inevitable and long as the people choose not to answer to answer particular or who provide inaccurate personal information on some issues are more or less evenly distributed throughout the study population or are distributed in a reasonably predictable way which allows researchers to introduce statistical controls for known sources of confounding, then when the dataset is interrogated by researchers for information on particular issues then it's likely that the degree of underreporting in both the chosen study group (e.g. people who have been diagnosed with a particular disease or condition since the study began) and the control/comparison group (people who haven't) will be sufficiently similar to cancel each other out when the data is analysed, not least because study participants cannot have been influenced in their decisions as to whether or not to disclose certain information either by the knowledge that they have been diagnosed with a particular condition or by the knowledge that researchers are investigating that condition and its possible causes and/or contributory risk factor.

In scientific terms both the study participants and, to a considerable degree, the researchers collecting the data, are 'blind' to the research question that that data will eventually be used to investigate and this limits the risk that underreporting of certain issues will bias the outcome of any future research.

In retrospective studies, however, both the researchers and study participants will be aware of exactly what the study is aiming to investigate before any questions about participants' medical history, etc. are asked – and, of course, the people in the study group will also know that they've been diagnosed at some point with the condition that the research is seeking to investigate – and this can create a problem if the people in either the study group or control group are influenced in

their decision as to whether to provide or withhold personally sensitive by what they know about either their own condition or the purpose of the research to a greater degree than the people in the other group.

So, any significant difference in the degree of underreporting of certain issues between the people in these two groups could potentially bias the results of the study to the extent where it fundamentally affects the study's result, particularly if the effect size associated with a particular issue or risk factor is relatively small or the degree of bias extremely large – and this is what researchers call either recall or response bias.

When it comes to ABC studies, the prevailing view amongst many epidemiologists is that women who have been diagnosed with breast cancer are, all other things being equal, more likely to overcome any reluctance they might experience when confronted with very sensitive personal questions, like 'have you had an abortion?' than women who haven't. This is because, in general, being diagnosed with a serious and potentially life-threatening condition tends to provide people with a very powerful incentive for them to cooperate with researchers as fully as possible, one that is personally felt to a much greater degree than is the case for people in the control group, who haven't been diagnosed with that condition, and where this occurs it will introduce a positive bias in the result of a study, i.e. the results will tend to show a stronger association between breast cancer and abortion than would be the case without this element of bias.

That said, the published evidence in this area is to date rather mixed, with some studies reporting a significant degree of response bias [e.g. Lindfors-Harris et al. 1991] and others finding only minor difference in reporting between study and control groups [e.g. Tang et al. 2000].

Brind, as you might expect, tackles the issue of recall/response bias in his own paper by trying to dismiss it as possible source of bias in his paper but the best that can be said for his argument is that it's all a bit of red herring. The many factors that could potentially exert an influence on response rates in ABC studies are too diverse and too variable across different countries, cultures and communities for estimates from any individual study to provide a reliable estimate of the possible impact of recall bias on another study conducted in a different location at a different time with a different study population.

So, if cross-cultural confounding, which is what we're talking about here, presents a barrier to investigating the potential impact of recall bias on retrospectively designed ABC studies then how do we know whether or not recall bias has had an impact on Brind's results and, if so, in which direction?

One way would be to conduct separate analyses of the data from prospective and retrospective studies and then compare the two. We know that studies conducted using a prospective design are much less prone to recall bias, so if recall bias is a confounding factor then this will become apparent by comparing the results of the two different types of study. If it isn't an issue that, all other things being equal, the results from prospective and retrospective studies should be near enough the same as makes no difference but if it is then the difference between the two will tell us not only whether recall bias is a confounding factor but also the direction of that bias, i.e. whether it leads to studies over or under estimating the scale of any statistical relationship between abortion and the prevalence of breast cancer.

Unfortunately, when it comes to Brind's paper there really isn't enough evidence included in it from prospective studies to make for a reliable comparison, not when you take into account all the uncertainties about study heterogeneity and the overall quality of the studies included in his

analysis, so the most we can conclude for now is that recall bias could be a confounding factor in Brind's study.

That leaves us with one final issue to discuss in relation to Brind's paper - effect size.

An odds ratio of 1.3 for the claimed association between induced abortion and an increased risk of breast cancer is, in epidemiological terms, a very weak effect size; far too weak, in fact, to sustain any strong claims of causality. By way of a useful comparison, epidemiological studies examining the prevalence of lung cancer in cigarette smokers compared to non-smokers typically generate odds ratios of anything from 20 upwards, with 35-40 being typical for countries such as the US and UK, leaving no room for doubt as to the causal nature of the link between smoking and lung cancer.

The best that can honestly be said for Brind's paper is that it suggests that there might be something in the ABC hypothesis that requires further investigation but given the weak effect size and the marked uncertainties about data quality, study heterogeneity and recall bias, claims that this study shows a clear causal link between induced abortion and a subsequent increased risk of breast cancer are wholly unsustainable. Brind's evidence simply isn't good enough, not matter what he might personally think.

I've spent quite a while working through Brind's paper and its flaws and limitations but for good reason because, if you've followed and understood everything so far then you should be well equipped to make sense of the other two papers we need to look at, the 2004 study published in *The Lancet* by Oxford University's Collaborative Group on Hormonal Factors in Breast Cancer (AKA Beral et al. 2004) and the new Chinese paper by Huang et al.

Beral et al. 2004.

So how does Beral et al. shape up in comparison to the earlier study by Brind et al.?

Well, to begin with it's a much larger study which takes in data from 53 epidemiological studies conducted across 16 countries giving a combined sample of 83,000 women who had been diagnosed with breast cancer. This includes the majority of, but not all, the papers included in Brind's study, eight of which were excluded from the main analysis because the researchers were unable to obtain access to the original data sets from those papers, this being part of the study's inclusion criteria.

They did, however, carry out a secondary analysis that included the results from these papers, which found that their exclusion would have no significant or appreciable impact on the results of the primary analysis:

Of the eight published studies that did not contribute at all to this collaboration only two had recorded information on abortion history prospectively. These two studies included 1516 women with breast cancer (only 3% as many women with breast cancer as in such studies that did contribute), and in them the combined relative risk for one or more induced abortions compared with none was 0.99 (95% CI 0.89–1.11). When the published results from these two studies were combined with the present results from the other studies with prospectively recorded information (upper part of figure 2), the overall relative risk was still 0.93 (0.89–0.97). For the studies with retrospective information not included here that had published relevant data (including published results from the one participating study that principal investigators requested be excluded from these analyses), the combined estimate of the relative risk of breast cancer associated with one or more reported induced abortions

was 1.39 (1.22–1.57); and when those published results were combined with the present results from the studies with retrospectively recorded information (lower part of figure 2), the overall relative risk was 1.14 (1.09–1.19).

By obtaining access to the original data sets for all the studies included in the main analysis, Beral et al. were able to stratify the sample data to ensure that women in each study were compared directly only with women in the same study and introduce a uniform range of statistical controls for known sources of confounding, reducing to a minimum the risk of confounding due to variations in study quality and study heterogeneity. Publication bias was addressed in part by issuing a call for unpublished data which resulted in the inclusion of data from three, at the time, unpublished studies, a Scottish prospective study and UK-based two retrospective studies.

So, even before we begin to look at any findings we can see that this study has adopted a much more rigorous methodology than that used by Brind, one which directly addresses some of that study's limitations. Crucially, where Brind's study included only three prospective studies which would have supplied at most around 20% of the data for his paper, thirteen of the fifty-three studies included in Beral et al. used a prospective design and these account for a little over 53% (44,000 women) of data in this paper. Beral et al. were, therefore, in a position to carry out separate analyses for studies that used prospective and retrospective designs and investigate the question of whether or not retrospectively designed ABC studies are prone to recall bias and, if so, in which direction that bias tends to operate.

So, time for some results and in the main analysis of the data from prospective studies, Beral et al. failed to find evidence to support claims of link between induced abortion and an increased risk of breast cancer in women:

For induced abortion, figure 2 shows the study-specific results and the combined results separately for studies with prospectively and retrospectively recorded information on abortion. Relative risks for the 13 studies with prospectively recorded information (upper part of figure 2) are close to, or slightly below, unity, and the weighted average of them yields an overall relative risk of 0.93 (95% CI 0.89–0.96; $p=0.0002$). Hence, neither the results from the individual studies nor their weighted average suggest any adverse effect on the subsequent risk of breast cancer for women with prospective records of having had one or more pregnancies that ended as an induced abortion, compared with women having no record of such a pregnancy. Furthermore, among the studies with prospective records of induced abortion, no significant variation in the results was found between those with objective and those with self-reported information (RR 0.93 [95% CI 0.88–0.97] and 0.92 [0.85–0.99], respectively; χ^2_1 for heterogeneity=0.04, $p=0.8$).

They did, however, find a substantial difference in the estimate of relative risk provided by prospectively and retrospectively designed studies which they concluded could not be accounted for by known differences in the women included in each type of study, leaving recall bias as the most likely cause and one which introduces a systematic positive bias consistent with the findings of Lindefors-Harris et al. 1991.

In short, if you look only at the evidence from prospectively designed studies, which are not prone to recall bias, then the evidence for an increased risk of breast cancer that Brind claimed to have found in women who had had an induced abortion, compared to those that hadn't, disappears completely.

Moreover, the ABC hypothesis predicts not only that we should see an increased risk of breast cancer in women who have had at least one induce abortion, it also predicts that we should see a

significantly higher risk in women who abort their first pregnancy rather than carry that pregnancy to term and that there should be a 'dose-response' effect as the number of abortions a woman has increases, i.e. more abortions equals a higher risk; and yet sub-analyses included in Beral et al. fail to support either of those predictions. For two or more abortions vs. one abortion the study reports and non-significant relative risk ratio of 0.96 while for abortion before or after the birth of the first child the relative risk ratio was statistically significant but was only 0.91, which could be taken as indication that aborting your first pregnancy actually reduces your long term risk of breast cancer compared to carrying that first pregnancy to term were it not for the fact that the effect size is far too small to support any such conclusion.

Beral et al. was, at the time of publication, and still is the largest, most comprehensive and most rigorous study of the relationship between induced abortion and subsequent risk of breast cancer and yet it produced no evidence whatsoever that would support any of the claims or predictions made by the ABC hypothesis while, at the same, providing fairly solid evidence to suggest that retrospectively designed ABC studies are not only prone to recall bias but also that the direction of that bias is positive and that these types of studies will, therefore, tend to inflated estimates of risk.

Beral et al. is not without its critics, notably Joel Brind who, since publishing his own paper, has confined his 'research' efforts in this field to firing off critical letters to various journals whenever an ABC study is published that doesn't support his findings which invariably claim that there a serious flaws in the newly published study which invalidate its findings. Brind, by the way, also appears to believe that the 'medical establishment' is lying to women and refusing to accept his findings as being the last word on the ABC hypothesis because of 'political correctness'.

Now I'm not going to address all the arguments that Brind levelled Beral et al. but I will pick on one which illustrates the extent to which his these arguments are rooted in his own personal biases and not in any kind of solid scientific reasoning. As I noted earlier, eight of the papers that Brind included in his study were excluded from this one because researchers were unable to obtain access to the original data, and three of those were old enough that none of the original authors could be contacted.

This prompted Brind to accuse Beral et al. of *selection bias*, i.e. of deliberately designing and carrying out their study in a way which excluded those papers because their findings support Brind's claim that induced abortion does lead to an increased risk of breast cancer, but as I've already noted, although these studies were excluded from the main analysis they were included in a secondary analysis which looked specifically at whether or not the inclusion of these papers would have had any effect on the study's results, and the answer was unequivocally 'No'. Including the relatively small amounts of data these papers might have supplied did increase the relative risk ratios in the main analysis slightly but by nothing like enough to alter the studies main findings at all, all of which the paper reports in its results section with the relevant descriptive statistics. Not only is there no selection bias there, but by failing to acknowledge that Beral et al. did clearly and openly address the question of whether the excluded papers would have affected their findings, Brind is resorting to dishonest argumentation in an effort to dismiss evidence which contradicts his own personal views.

To wrap things up on Beral et al. I should also note that since its publication in 2004 two further large scale longitudinal prospective studies; Michels et al. (2007), which uses data from the Nurses' Health Study and DeLellis Henderson et al. (2008), which takes its data from the California Teachers Study, have published findings on induced abortion and breast cancer and, like Beral et al. neither found evidence of an increased risk of breast cancer in women who have had an abortion.

Finally, then, the way is clear for us to consider the new Chinese study by Huang et al.

Huang et al. 2013.

Okay, so basic details first and what we have here is a meta-analysis of data from 36 studies (2 cohort, 34 case control) conducted across 14 provinces of China with a combined sample size of 41,690 women, including controls. So it's a larger study than Brind et al. but not as large as Beral et al. and all the data comes from a single country.

Superficially, the methodology here more closely resembles that of Brind than Beral. The majority of studies included in the analysis use a retrospective design and although in some cases Huang was able to obtain access to the original data sets on which studies were based, the bulk of their data still comes solely from published papers. 14 studies reported findings solely for induced abortion, the rest for induced and spontaneous abortion (i.e. miscarriage) which is something of weakness but not necessarily an insurmountable one.

That said there are also several obvious improvements on Brind study.

Huang's paper does include a systematic analysis of study quality which uses the recently developed [Newcastle-Ottawa scale](#), a grading system designed specifically for dealing with non-randomised cohort and case control studies, on the basis of which 8 of the studies were given the highest quality grade 'A', 26 were graded 'B' and just 2 were given the lowest grade, which is 'C'.

There is a statistical investigation for publication bias which indicates that this is not a significant issue for this paper and there is also a statistical investigation of study heterogeneity, which is rather high at $I^2=88\%$ for the induced abortion studies and $I^2=71.9\%$ for those which provided combined for both induced and spontaneous abortions, which suggests there maybe uncontrolled sources of confounding that may affect the results. That said, there is also a secondary analysis for study heterogeneity, which is presented separately to the main paper and which covers only the 18 papers which fell within the standard error range in the publication bias funnel plot which found no heterogeneity in these studies, so we appear to have a sub group of studies at the core of Huang's paper which are all very similar in design, study population, etc. all of which present findings that are broadly in keeping with the headline findings of the study as a whole.

And the main findings?

Huang did find a statistically significant association between abortion and an increased risk of breast cancer in women with one recorded induced abortion (OR = 1.44, 95 % CI 1.29–1.59) and that this risk increased with an increase in the number of abortions; for 2 abortions the odds ratio is 1.76 (95 % CI 1.39–2.22) and for three abortions rises again to 1.89 (95 % CI 1.40–2.55).

On the face of it, this appears to support Brind's original findings and flatly contradicts those of Beral not just in terms of finding an association between abortion and breast cancer but also in finding evidence of the dose-response effect predicted by the ABC hypothesis.

So, that puts the ABC hypothesis back on the agenda, yes?

Well that is certainly what the anti-abortion lobby thinks as it has now been reported that LIFE has written directly to the Parliamentary Under Secretary of State for Public Health, Jane Ellison, asking her to conduct an inquiry into the alleged link between abortion and breast cancer on the back of this particular study.

However, all is not what it might appear when you start to look at the detail of Huang's paper as there are one or two problems which raise serious questions about its findings.

I'll start with a fairly trivial issue just because it's one I personally found rather amusing and it is to be found in this section of the paper's introduction:

Although the two reviews mentioned above [7, 8] had focused on the association between IA and breast cancer, they did not include several important studies, such as the study from Jiangsu [11]. Omission of these important studies undoubtedly biased the summary results.

The two review referenced here as 7 & 8 are the studies by Brind et al. and Beral et al. which were published in 1996 and 2004 respectively. The Jiangsu study, the omission of which 'undoubtedly biased' the findings of the other two studies was not published until 2012. It is actually no more than a relatively small retrospective case control study (669 cases, 682 controls) which didn't even begin recruiting subjects until June 2004. On the face of it, Huang has invented a whole new class of research bias, temporal confounding, which arises as a result of researchers being unable to travel into the future to obtain data from studies that haven't even commenced at the point at which their own work is published.

Okay, so this could be a case of something getting garbled in translation but one still has to be a little careful when encountering such statements in Chinese research as there are still parts of the Chinese Academy that are prone to exaggeration for largely cultural and political reasons, a tendency rooted in the country's longstanding belief in its own exceptionalism which, if anything, Communist rule in the post-war period served only to exacerbate.

Having looked closed at the Jiangsu study there appears to be nothing to commend it as being of particular importance over and above most of the other case control studies included in Huang's paper, or indeed over similar studies in either Brind or Beral, beyond it being one of the eight studies that was given a 'A' grade in Huang's study quality review but that, in itself, leads directly to another area of concern.

As already noted, Huang made use of the recently developed Newcastle-Ottawa scale to systematically grade the study quality of paper included in analysis presenting the results of this review in a table listing each of the papers with a grade awarded on a 3-point scale, either 'A', for the highest qualities studies, 'B' or 'C'.

However, if you look at the core documentation for the Newcastle-Ottawa scale what you will find is that it uses a 'star' system under which studies are assessed in terms of three broad perspectives; the selection of the study groups; the comparability of the groups; and the ascertainment of either the exposure or outcome of interest for the type of study, the intention being that when reporting these grades researchers should individually report the star rating for each of these three domains – this can be up to 4 stars for study group selection, 2 for comparability and 3 for exposure/outcome on interest.

If used and presented in the recommended manner, the Newcastle-Ottawa scale should provide a fairly clear overview of the potential strengths and limitations of not just individual studies but also across the board for all studies included in a review paper or meta-analysis. One could, for example, very quickly ascertain whether or not studies included in a review exhibited any obvious similarities in either their strengths or limitations which in turn may highlight potential issues with a meta-analysis' findings, i.e. if a large number of studies score poorly on study group selection then there may be uncontrolled sources of confounding in the data and would therefore have to look closely at

what, if any, statistical controls were introduced into the meta-analysis to control or correct this problem.

Huang's presentation of his study ratings is, however, distinctly non-standard. First he converts the star ratings awarded to each study to a single numeric rating (0-9) by simply adding together the total number of stars and then he converts this numeric rating into the a three point scale ('A', 'B' or 'C') using uneven scale boundaries, 0-4 for 'C', 5-7 for 'B' and 8-9 for 'A'. Although the paper explains how these grades were arrived at, what it doesn't even try to explain is why a non-standard presentation method has been used nor whether this presentation method has been subjected to any kind of testing to validate the 3-point scale that is used. Huang may well have assess the study quality of the papers included in his analysis but in presenting that information to the reader he has not only stripped his evaluation of all the clarity built into the Newcastle-Ottawa scale by its designers but has also given his readers a derived grading schema that cannot be accurately interpreted without going to the time and trouble of digging out copies of 36 papers and repeating the grading exercise.

That I find a little troubling. If you are going to make the effort to systematically evaluate the quality of the studies included in your analysis why then sabotage your own efforts by presenting the results of you evaluation in a non-standard and seemingly entirely arbitrary manner which serves to obfuscate rather than clarify your findings?

It is an approach that makes no sense at all unless, for some reason, you don't want to open up your evaluation to independent scrutiny.

Another concern, which is not immediately apparent unless you take the time to chew over Huang's paper in detail, is the inclusion of contradictory statements and assertions in relation to prevailing social attitudes to abortion in China.

In the introduction section we are told that:

As one of the countries with the highest prevalence of IA, in China, it is particularly important to clarify the association between IA and breast cancer risk. The lack of social stigma associated with IA in China may limit the amount of underreporting and present a more accurate picture of this association [10].

The reference here is to a paper by Sanderson et al. published in 2001 which reports on a cohort study conducted in Shanghai which failed to find evidence for a link between abortion and an increased risk of breast cancer in the study population the abstract to which begins with this statement:

Studies of the association between induced abortion and breast cancer risk have been inconsistent, perhaps due to underreporting of abortions. Induced abortion is a well-accepted family planning procedure in China, and women who have several induced abortions do not feel stigmatized.

There is, however, an assumption here that the majority of abortions in China result from the State's one-child policy, i.e. that most women undergoing an abortion have already given birth and are using the procedure to limit the size of their family and, indeed, it was on the basis of this assumption, which could easily be valid for Shanghai but not for other parts of China, that Joel Brind dismissed the negative findings of this paper, arguing that these were inapplicable to abortion in the United States where the majority of abortions are carried out to delay starting a family. That said, Brind has also very noticeably chosen not to apply the same argument to Huang's study, which he's

been actively promoting of late, which could be taken as indication that he's forgotten what he had to about Sanderson's earlier paper but in reality is more a validation of Polly Newcomb's observation that Brind's personal biases render him incapable of critically evaluating the evidence from ABC studies.

However, in the discussion section Huang states that:

Our results might be confounded by additional factors. First, some abortions performed before marriage might be included. However, these abortions were very few, and probably would not be reported in China [9], as they are less socially acceptable and are associated with more stigmas.

So there isn't actually a complete lack of social stigma attached to abortion after all, particularly where it takes place before marriage and is associated with 'more stigmas' – again the point needs to be made that in a country the size of China the extent to which abortion may be stigmatised can and almost certainly will vary considerably between different provinces and between urban and rural areas just as it does in the United States where one would fully expect to find very different social attitudes towards abortion in, say, San Francisco than would be found in a small town in Mississippi.

However, we are again assured that abortion prior to marriage in China is extremely rare and unlikely to be a source of confounding – and if you haven't already figured this out for yourself, both statement are included in the paper to support an argument that recall bias is not a relevant issue in ABC studies conducted in China which, if true, would be a very important argument for Huang to make successfully given that the vast bulk of his data comes from retrospective studies.

However, contrast the views given in those statements with this [2009 news report from the China Daily](#), which is the largest state-owned English language newspaper published in China:

Inadequate knowledge about contraception is a major factor in the 13 million abortions performed in China every year, research shows.

This is an unfortunate - and avoidable - situation, experts said, and improvements need to be made.

Li Ying, a professor at Peking University, said Wednesday that young people need more knowledge about sex.

...

Government statistics show that about 62 percent of the women who have abortions are between 20 and 29 years old, and most are single.

Wu said the real number of abortions is much higher than reported, because the figures are collected only from registered medical institutions.

Many abortions, Wu said, are performed in unregistered clinics.

Also, about 10 million abortion-inducing pills, used in hospitals for early-stage abortions, are sold every year in the country, she said.

'Wu', by the way, is named in the article as Wu Shangchun, a division director of the National Population and Family Planning Commission's technology research centre, so we can safely take this report to be the Chinese state communicating its official view through a state-owned newspaper

based on its own research and official statistics, and what the Chinese state is saying here flatly contradicts Huang's assertion that pre-marital abortions are rare.

The official line here from the Chinese state is that there are, according to its best estimates, around 13 million abortions performed in China to go with around 20 million live births a year, so on those figures around 40% of all pregnancies end in abortion, but there are also seemingly substantial uncertainties around unregistered clinics and the private sale of medical abortifacients which could mean that even that 13 million figure is an underestimate.

And, of course, it states that a clear majority of women having abortions are under 30, most of whom are single.

China clearly seems to have a few problems not least amongst which are a couple which you may find eerily familiar:

Sun Xiaohong, of the educational center of Shanghai's Population and Family Planning Commission, said she found it difficult to promote sex education in schools because some teachers and parents believe it will encourage sex.

Sun Aijun, a leading gynecologist at Beijing Union Hospital, said there also is a misconception among some women that the contraceptive pill is unsafe.

Mmm... I wonder where those misconceptions may be coming from? *innocent face*

On the face of it, Huang's argument that Chinese ABC studies can be considered to be free of any confounding from underreporting or recall bias may very well not fly in the face of what the Chinese government is now saying about abortion in China, not without further evidence to show that the trend towards young unmarried women having abortions is indeed a relatively recent innovation in the country; too recent in fact for it to be a relevant factor in ABC studies of older women. China did, after all, introduce its controversial 'one-child policy' in 1979 and was also amongst the early adopters of Mifepristone, the main drug used to perform medical abortion, which was approved for use in China to end pregnancies up to 49 days gestation as far back as 1986, and so one cannot simply assume that it is only recently that single women have become the primary users of abortion services in the country not least because of the effect that a child born outside marriage could have on a young woman's future marriage prospects.

This, of course, prompts an obvious question: what actual evidence is there to back up this assertion?

The citation on which Huang places most reliance is a 2001 paper by Sanderson et al. which reports findings from the Shanghai Breast Cancer study. This is one of the larger case control studies included in Huang's analysis but still provides data only on 2844 women living in one location in China. Nevertheless, Sanderson includes the following argument in his discussion of his findings:

Underreporting of induced abortions is unlikely in our study given its' widespread use in China as a family planning method in case of contraceptive nonuse or failure. [21] China has had a series of family planning campaigns in place since 1956. Induced abortion was legalized in China in 1957 around the time most of the women in this study were beginning their childbearing years. [6] The procedure is free of charge and readily available. Because the primary method of family planning in China at the time most women in this study were using contraception was the intrauterine device that was known to have high failure rates and women were expected to have a child soon after marriage, women oftentimes had

more than 1 abortion after the birth of their first or second child but not before their first live birth. Because of this and because Chinese women who have several induced abortions do not feel stigmatized, we believe that the information on abortion collected in our study is rather accurate.

But this is almost entirely based on a priori reasoning from both his observations of the women in his own study in which, incidentally, just over 9.5% reported that their first abortion took place before their first live birth, and on the ready availability of abortion in China, all of which makes the leap necessary to get from this small group of women to all Chinese women who have had an abortion since 1957 very large indeed.

Sanderson does include a reference (21) to a 1990 paper by Li et al. which looks at the general characteristics of women in China who were having an abortion at that time but, having tracked down the abstract for that paper it turns out that this was a self-report study conducted in 1985 of just 1,200 women living in just two provinces (Szechuan & Jiangsu) plus the municipality of Shanghai in which the main focus was more on contraceptive use at the time that women fell pregnant in circumstance in which they found it necessary to have an abortion and study abstract includes an extremely important caveat:

The data presented here are limited and cannot be generalized to the larger population.

That said, the information that follows this caveat seems fairly consistent with the concerns that China Daily were reporting almost 25 years later:

However, they do shed some light on the contraception characteristics of a group of women who undergo abortion procedures in China. Their response to questions to contraceptive behavior prior to abortion suggests that the problem, in part, is behavioral. For example after the expulsion of the IUD, no other method was substituted to avert pregnancy. In order to alleviate the problem of contraceptive failure, and subsequent abortion, there are policy as well as training and education implications for the state.

So what we have here are some very strong assertions as to how abortion is used in China and how this differs, almost fundamentally, from its use in the UK and other Western countries but very little evidence to support those assertions; and what little evidence there is seems highly unlikely to be generalizable to the wider population. And yet, on the back of these assertions, Huang more or less completely dismisses recall bias as a relevant issue in Chinese ABC studies without any further investigation.

For me this is just not a sustainable position, particularly when it's apparent that even the Chinese state is lacking any clear or comprehensive view of not only how abortion is used but even how many abortions are actually taking place, and that casts doubts not only over Huang's analysis but over pretty much every ABC study conducted in China.

Bearing all that in mind, I've left by the most serious problem with Huang's paper to last, in part because it gets a little technical in places but also because I like to end on a bit of a bombshell.

Notwithstanding the headline findings of his analysis, which appears to lend support to the ABNC hypothesis both in terms of indicating an increased risk of breast cancer in women who have had abortions and a dose-response effect in which this risk increases as the number of abortions increases, Huang still has a significant problem to overcome which stems from his use of the Newcastle-Ottawa scale to grade the quality of the studies included in his paper.

This problem arises because of the eight studies given the highest 'A' grading, which included the two cohort studies, six produced negative results which showed either no increase in breast cancer risk in the abortion group or a very small but non-significant increase. Given that only 27% of the data included in the paper comes from 'A' grade studies, with only 21% coming from those with negative findings, the obvious criticism to level at Huang's study is that the best quality evidence he has is effectively being swamped by the inclusion of a large number of lower quality studies where there is likely to be a much greater risk of confounding and, especially, given the reliance on retrospective studies, positive bias arising from uncontrolled recall bias.

Huang attempts to pre-empt this line of argument in the following section of the paper:

Since the positive association between IA and incident breast cancer was first presented by Segi et al. in [58], several studies supported this association [59–62]. However, some other studies, including two important studies from Shanghai [9, 10], found a null or similar association. Inadequate choices of the reference group might be one of the most important determinants of the different results. In fact, the prevalence of IA in the control group were more than 50 % among both the two Shanghai studies (51 % in Ye et al. [9], and 66 % in Sanderson et al. [10]), and among several other included studies with NOS of 8–9 (80.4 % in Qiu et al. [31], 68.3 % in Zhang [25], 63.0 % in Wang et al. [32], and 62.7 % in Wang [27]). As argued by Brind and Chinchilli [14], once the prevalence of a given exposure rises to a level of predominance in the control group, statistical adjustment cannot remove all the confounding caused by the adjustment terms. This was well exemplified by the meta-regression analysis in our study (Fig. 6). It was also the main reason why we did not observe an increased risk of breast cancer in the subgroup analysis based on Shanghai studies, studies with a NOS score of 8–9, and cohort studies, because both studies of Sanderson and Ye were conducted in Shanghai [9, 10] and with a NOS score of 8–9, and the study of Ye was one of the two cohort studies.

And there is indeed a graph (Fig 6.) which appears to show that studies with a higher incidence of induced abortion in the control group tend to exhibit lower odds ratios and therefore either a lower or no association between induced abortion and breast cancer risk.

So, what Huang is seeking to demonstrate here is that because most of the higher quality studies are also studies in which there is a relatively high prevalence of induced abortion in the control group, these studies are actually less reliable than the lower grade studies in his paper and in this Huang is entirely reliant on an argument advanced by Joel Brind and his, at the time, main collaborator Vernon Chinchilli in [a comment published in the British Journal of Cancer](#) in 2004 in response to a 2002 Chinese ABC study by Ye et al. which failed to find evidence of an increased risk of breast cancer associated with induced abortion.

This is reference 14 in Huang's paper and the relevant section of Brind & Chinchilli's argument is this one:

Another important difference, however, between these Chinese study populations and those of most western industrialised countries, is the very high prevalence of induced abortion in China. In the study of Ye et al, the prevalence of induced abortion is 51%, and in the study of Sanderson et al, it is 66%. The validity of any observed association – null or otherwise – between a given exposure and a given disease outcome, rests upon, among other things, the unexposed population's serving as a typical, appropriate reference group. Once the prevalence of a given exposure rises to a level of predominance, it is prudent to

ask whether indeed the unexposed comparison group has instead become a subgroup, which is unexposed for some reason that bears relevance to its risk profile for the disease in question. In such a case, statistical adjustment cannot remove all such confounding, since the calculation of the adjustment term will necessarily be underestimated. In the case of the Shanghai study population, the confounding by parity and age at first birth would not be fully corrected for, and the relative risk for induced abortion would remain underestimated.

So what Brind and Chinchilli are saying here is that Ye's paper, and by extension any study where there is high prevalence of induced abortion in the control group, is likely to be subject to a degree of uncontrolled residual confounding which will cause it to underestimate the scale of any association between induced abortion and breast cancer, and Huang's meta-regression analysis would appear to support this view if, *and only if*, Brind and Chinchilli are correct. If, however, Brind and Chinchilli are incorrect then Huang's meta-regression will show that the opposite is true and that studies in which there is a relatively high prevalence of induced abortion in the control group, which would be consistent with the high overall prevalence of abortion in China where, as we've already seen, official estimates suggest that at least 40% of pregnancies end in abortion, are those which provide a more accurate estimate of the relationship between abortion and breast cancer, in which case his attempt to negate any criticisms arising from the fact that most of the best quality studies included in his paper produced negative results will collapse under the weight of its own faulty assumptions.

And here is where Huang runs into a very serious problem because Brind and Chinchilli's comment in the British Journal of Cancer [attracted a reply from the two of the co-authors on Ye et al, DB Thomas and RM Ray](#), which was published in the same edition of the journal, immediately after that of Brind and Chinchilli:

In their letter with regard to our paper on induced abortions and breast cancer, Brind and Chinchilli essentially suggest that residual confounding by age at first birth and parity may have caused us to underestimate the odds ratio (OR) for breast cancer in relation to induced abortion. We disagree. In paragraph 2 of their letter, they suggest that women in China who did not have an induced abortion would be more likely to be nulliparous and to have had their children later in life than women who had an abortion, that the women unexposed to abortions were therefore at higher risk of breast cancer than those with an abortion, and that the true OR in relation to induced abortion was thus underestimated. This is not correct. Few women in our study cohort were nulliparous and, as stated in our paper, the results were virtually unchanged when the analyses were restricted to gravid or parous women. Because of the one child per family policy in China, which became operational in the early 1980s, older women in our study tended to have larger numbers of children than younger women, and to have begun child bearing at an earlier age. Because of this, after controlling for age, the number of children was not a confounder, and age at first birth was only a weak confounder. During the time period covered by our study, abortions were almost always performed to limit family size. The decision to have an abortion would thus have been made after the birth of one's first child. Therefore, age at first birth would not necessarily be earlier for women with an abortion than for women of the same age without an abortion, as Brind and Chinchilli contend.

Brind and Chinchilli point out that our crude OR for breast cancer in relation to induced abortion is 0.93, and our OR adjusted for age and age at first birth is 1.06. In the next paragraph, they suggest that confounding by parity and age at first birth would somehow not be fully controlled for by adjustment because of the high prevalence of induced abortion

(51%) in our study population, and therefore that the OR of 1.06 should actually be higher. We fail to understand how the prevalence of the exposure could directly influence the confounding effect of other factors. If one were to (hypothetically) conduct a randomised trial of abortion and breast cancer, the most efficient design would be to assign 50% of the women to an abortion group and 50% to a control group. Confounding would be controlled for by the randomisation. In our study, confounding was essentially controlled for by stratification on the potentially confounding variables of concern. With a prevalence of exposure to abortions of 51%, we have the optimal power to detect a true association and to control for confounding. If the prevalence of abortions in the population were closer to 0%, it would have been more difficult to control for confounding by stratification because of smaller numbers in the exposed group.

You cannot, I think, get a more straightforward answer than 'This is not correct' and yet of Thomas and Ray's response to Brind and Chinchilli's argument there is not a single mention in Huang.

That, for me, is a serious problem. Back in my own university days one thing that was drilled into me was that, when writing up research in which you are relying at any point on a reference or argument that is subject to dispute, the very least you should do is note the existence of that dispute and provide references sufficient for anyone reading your work to investigate the issue for themselves. Ideally, of course, your own research should at least in part address the disputed point and, if possible, supply evidence to support your decision to rely on a particular position but if that's not an option that it is at least good practice to state your own reasons for favouring one view of an issue or one position over another.

Huang does none of this, he simply presents Brind and Chinchilli's argument as a *fait accompli* and carries on to try and knock over the evidence he has from the best quality studies included in his paper, most of which contradicts his headline findings.

Although one cannot completely rule out the possibility that this is due to an oversight or simply poor scholarship it is nevertheless difficult to see quite how Huang can have failed to run across Thomas and Ray's rebuttal of Brind and Chinchilli's argument when reviewing the relevant literature for his own paper, particular when the two articles appear to have been published on consecutive pages of the same edition of the same journal. What this looks like, regrettably, is cherry-picking, or to put it more formal terms; researcher (or experimenter) bias, although if this indeed the case what we cannot be sure of is whether it stems simply from the author(s) desire to produce an eye-catching piece of research in order to maximise its publication/citation potential or whether there is some underlying agenda at work.

So, all things considered, what can we actually say about Huang's study?

Well, although it does appear to make several improvements on Brind's 1996 study, which overall it most closely resembles in terms of both its basic methodology and, of course, its headline findings these improvements are, by and large, mostly cosmetic and rather superficial.

On the key criticisms levelled at Brind's paper, that the overall effect size reported was too weak to sustain any kind of strong claims for a causal link between induced abortion and a subsequent increased risk of breast cancer in women and that, in all probability, even that effect was likely to be due mostly, if not entirely, to the confounding effects of recall bias, Huang's paper offers no significant improvements. Indeed, Huang's decision to obfuscate the results of his systematic review of study quality by using a non-standard presentation method and a seemingly arbitrary and unvalidated derived 3-point scale coupled with his reliance on Brind and Chinchilli's dubiously

premised argument and comprehensively rebutted argument and, of course, his failure to acknowledge even the existence of that rebuttal, in order devalue the evidence supplied by the highest quality papers included in his analysis rather tends to suggest that Huang and/or his co-authors may be well aware of this study's serious weaknesses, which they have gone out of their way to try and conceal.

Conclusion

So, do the findings here either call into serious question or overturn those reported by the Collaborative Group on Hormonal Factors in Breast Cancer (Beral et al.) in 2004?

No, absolutely not.

For the time being, at least, Beral et al. remains the definitive study in the field by virtue of superior and far more rigorous design and the breadth of the data it incorporates into its analysis, particularly in terms of the data accrued from prospective studies where recall bias is highly unlikely to influence the outcome of such studies and introduce bias into their findings.

Consequently, any suggestion that the findings reported in Huang et al. 2013 have put the ABC hypothesis back into play are, to quote Bentham, 'nonsense on stilts'.

A Note on Agenda Driven Research.

One thing that inevitably makes the newly published study by Huang et al. particularly attractive to the anti-abortion lobby, other than its relative size, is its source, i.e. China.

It has been well understood, for some considerable time, that the scientific literature on abortion emanating from the United States has been rather tainted by the efforts of a number of prohibitionist researchers, including Joel Brind, Vincent Rue, Priscilla Coleman and David S Reardon, to construct a body of literature in support of their religious/moral agenda on abortion, efforts which to some extent run aground after one or two key figures were caught playing fast and loose either with their credentials or their research practices.

David Reardon, for example, more or less stopped putting his name to any kind of published research or correspondence with scientific journals after he was 'outed' by Chris Mooney, in 2004, as having obtained his 'PhD' from a non-accredited 'diploma mill' while Priscilla Coleman was only relatively [found by researchers at the Guttmacher Institute](#) to have made [a number of false statements and to have drawn erroneous conclusions](#) in an abortion-mental health study published in the Journal of Psychiatric Research in 2009 and in a corrigendum to that same paper issued in 2011.

As for Joel Brind, although his academic reputation in his own field (endocrinology) remains intact, he has found it increasingly difficult to get his criticisms of ABC studies which generate findings that contradict his personal belief that there is a link between abortion and breast cancer into print. In the case of the two large prospective studies published in the United States since Beral et al. (Michels et al. 2007 and DeLellis Henderson et al. 2008) Brind was forced to resort to publishing in JPandS (Journal of American Physicians and Surgeons) which is not only not indexed by PubMed but is also widely regarded as having much the same relationship to a genuine scientific journal, like Nature or the New England Journal of Medicine that Conservapedia has to the Encyclopaedia Britannica.

Research originating in China which appears, at face value, to support claims that induced abortion gives rise to an increased risk of breast cancer is doubly attractive to the anti-abortion because it not

appears to support their agenda but, crucially, it derives from a country where it would appear that researchers would be actively pursuing any kind of religious or moral agenda that might influence their research practices or findings.

But is that necessarily true?

Because there seems to be no obvious religious or moral agenda, it [does not automatically follow](#) that there is no agenda at all:

Local authorities in China this week announced that Zhang Yimou, the acclaimed director of such films as *House of Flying Daggers* and *Raise the Red Lantern* and the brain child behind China's Olympic ceremonies in 2008, would be fined 7.5m yuan (£750,000) for violating China's one-child-policy. Admitting to having two sons and a daughter, Zhang explained that it was his "traditional way of thinking" to aspire for a bigger family that led to his transgressions with the law.

Traditional Chinese culture, steeped in Confucian principles, sees family units as the foundation of society and prosperity within the family forms a key part of this – which often means having multiple children (with an emphasis on male heirs) to support parents and grandparents. As a result, the one-child-policy, introduced in the late 1970s, is in fact an affront to traditional beliefs. While economic growth in China has manifested in an improved quality of life and a ravaging consumerist culture, couples and grandparents look to find true meaning in the bliss of a different prosperity: the children of the next generation.

Cutting through the high-minded bullshit, what this all boils down to is that China's rapidly growing middle class has been taking a good long look at its collective bank balance and is thinking to itself, "Why is state telling us we can only have one child when the money's rolling in and we can easily afford to raise two or three?". Traditional values and spiritual epiphanies be damned, we're minted and we don't see why we should face these kind of restrictions – and to be fair, they have a point because the one-child policy was originally a response to fears that the Chinese population was rapidly heading towards outgrowing the country's ability grow enough food to feed its citizens, or perhaps comrades, whatever the correct term might be.

Enforced population control is one of those idea that a lot of people will accept, however reluctantly, if the alternative appears to be starvation, but as China become increasing wealthy and the threat of starvation rapidly recedes, at least for the middle classes, then that argument will inevitably begin to look less and less convincing and the one-child policy, which only actually applies to around 36% of the Chinese population once you factor in the exemptions for rural families where the first child is either female or has a disability and those for ethnic minorities and the exemptions for residents of Macao and Hong Kong, will increasingly begin to look outdated and unnecessarily oppressive.

And under those circumstances, evidence which appears to indicate a connection between abortion, which is commonly used to limit family size, and an increased risk of breast cancer, which is very much on the rise in China, will inevitably look like a potentially useful lever in the argument against imposing strict limits on family size, especially when you consider that the people doing the research are also very likely to be amongst the same people who've been looking at their bank balances and thinking, "Mmm... you know we could easily afford more kids were it not for the fact that the state would take away my job if we did have any more".

We cannot, of course, be sure whether or not this kind of thinking might have come into play when Huang and his co-authors were putting together their analysis, but it is very much a possibility and it

does go to show that just because the kind of anti-abortion agenda were used to here in the UK or the United States appears to be conspicuous by its absence in countries like China, it doesn't mean that Chinese research is automatically going to be agenda free.

REFERENCES

Baer, H. J., Collins, L. C., Connolly, J. L., Colditz, G. A., Schnitt, S. J. and Tamimi, R. M. (2009), [Lobule type and subsequent breast cancer risk: Results from the Nurses' Health Studies](#). *Cancer*, 115: 1404–1411. doi: 10.1002/cncr.24167

Brind J, Chinchilli VM, Severs WB, Summy-Long J. [Induced abortion as an independent risk factor for breast cancer: a comprehensive review and meta-analysis](#). *Journal of Epidemiology and Community Health* 1996;50:481-496.

Brind J, Chinchilli VM (2004) [Breast cancer and induced abortions in China](#). *Br J Cancer* 90: 2244–2245

Collaborative Group on Hormonal Factors in Breast Cancer [Beral et al.]. [Breast cancer and abortion: Collaborative reanalysis of data from 53 epidemiological studies, including 83,000 women with breast cancer from 16 countries](#). *The Lancet*. 2004;363:1007-1016.

Daling JR, Malone KE, Voigt LF, White E, Weiss NS. [Risk of breast cancer among young women: relationship to induced abortion](#). *J Natl Cancer Inst* 1994; 86:1584-92.

DeLellis Henderson K, Sullivan-Halley J, Reynolds P, Horn-Ross P, Clarke C, et al. [Incomplete pregnancy is not associated with breast cancer risk: the California Teachers Study](#). *Contraception* 2008;77:391-396.

Huang Y, Zhang X, Li W, Song F, Dai H, Wang J, Gao Y, Liu X, Chen C, Yan Y, Wang Y, Chen K. [A meta-analysis of the association between induced abortion and breast cancer risk among Chinese females](#). *Cancer Causes & Control*. November 2013.

Li VC, Wong GC, Qiu S-H, Cao F-M, Li P-Q, Sun J-H. [Characteristics of women having abortion in China](#). *Soc Sci Med* 1990;31: 445–53.

Lindfors-Harris B-M, Eklund G, Adami H-O, Meirik O. [Response bias in a case-control study: Analysis utilizing comparative data concerning legal abortions from two independent Swedish studies](#). *Am J Epidemiol* 1991; 134: 1003-8.

Michels K, Xue Fei, Colditz G., Willett W. [Induced and Spontaneous Abortion and Incidence of Breast Cancer Among Young Women](#). *Arch Int Med* 2007 167:814-820.

Michels KB, Willett WC. [Does induced or spontaneous abortion affect the risk of breast cancer?](#) *Epidemiology*. 1996 Sep;7(5):521-8.

Ramakrishnan R, Gann PH, Wiley EL, Khurana KK, Khan SA. (2004). [Normal breast lobular architecture in breast biopsy samples from breast cancer cases and benign disease controls](#). *Breast Cancer Research and Treatment*. 86(3):259-68.

Sanderson M, Shu XO, Jin F et al (2001) [Abortion history and breast cancer risk: results from the Shanghai breast cancer study](#). *Int J Cancer* 92:899–905

Tang MT, Weiss NS, Daling JR, Malone KE. [Case-control differences in the reliability of reporting a history of induced abortion](#). *Am. J. Epidemiol.* 2000;151 (12): 1139–43.

Thomas DB & Ray RM (2004). [Reply: Breast cancer and induced abortions in China](#). *Br J Cancer* 90:2245–2246

Ye Z, Gao DL, Qin Q et al (2002) [Breast cancer in relation to induced abortions in a cohort of Chinese women](#). *Br J Cancer* 87:977–981