Post-Abortion Mental Health Research: Distilling Quality Evidence from a Politicized Professional Literature

Priscilla K. Coleman, Ph.D.

Elective abortion is unique in medicine for two key reasons. First, the 1973 Supreme Court decisions of Roe v. Wade and Doe v. Bolton declared the procedure to be a woman’s constitutional right; no other medical procedure has been similarly designated. Second, elective abortion results not only in an unnatural reversal of a woman’s body to a non-pregnant state, but also in the death of a biologically distinct, most often normally developing human embryo or fetus. Given differing perspectives on these issues, evaluation of whether or not abortion may initiate or exacerbate mental health problems is often influenced by the philosophical, moral, and political worldviews of the researchers involved.

Abortion is common in the U.S., with approximately 30 percent of women experiencing at least one during their reproductive years, underscoring the need for serious research examining abortion as a risk factor for psychological distress. The objectives of this article are to review how risk factors are established, provide evidence of bias in professional publications, and illustrate the use of a methodologically based rubric to raise the bar in future research. Women deserve and have the right to base medical decisions on empirically sound information interpreted in an ethically balanced and reasoned manner. Use of a standardized set of criteria is essential to interpreting individual study results, and is paramount in ranking studies incorporated into narrative and quantitative reviews. Without agreement on basic methodology necessary to assure reliable and valid data, and adherence to a complete set of criteria, the choice of standards to apply can be used to manipulate conclusions.

Generally speaking, a risk factor is any variable established to increase the probability of a particular outcome. Risk factor data are used in medicine and psychology to understand etiology, warn patients of potential problems associated with various exposures, and develop effective prevention and intervention protocols to maximize health. Assessment of degree of risk is expressed in terms of absolute risk, which relates to the chance of developing an illness over time (e.g., a 10 percent lifetime risk of suicide), or in terms of relative risk, which is a comparison of the probability of an adverse outcome in two groups (e.g., abortion vs. childbirth). Most of the research conducted on abortion and mental health concerns relative risk.

As the evidence for a risk factor accrues over time, scientists gain confidence that a particular exposure independently contributes to specified outcomes; however, the results of any individual study cannot be interpreted as definitive evidence. In the absence of the ability to manipulate reproductive outcomes, exploration of abortion as a risk factor for mental health problems requires a careful methodological analysis of the published literature. Specifically, the integrity of each study, and ultimately the broader establishment of abortion as a risk factor and possible causal agent in mental health problems, is based on the extent to which studies accomplish particular ends in line with the Bradford Hill (BH) criteria. First, abortion should be demonstrated to precede the mental health problem, accomplished with prospective or longitudinal data collection (BH, temporality). Second, differences in abortion history (abortion, no abortion) must be systematically associated with differences in mental health status (BH, association or co-variation). Third, plausible alternative explanations for significant correlations between abortion and mental health must be ruled out using established measures for controlling extraneous variables (e.g., by employing a case control design or statistically removing the effects of potential confounds from the outcome measure) (BH, specificity). Fourth, further evidence is derived when studies demonstrate a “dose effect” wherein the number of abortions is shown to be positively correlated with the seriousness of impairment (BH, biological gradient or dose-response).

Finally, when scientists are able to provide evidence for mediating mechanisms (e.g., feelings of loss, guilt, shame, fears regarding future pregnancies, feelings of partner abandonment, etc.), contributing theoretical underpinnings for abortion-related declines in mental health, there is further substantiation (BH, plausibility). The consistency and magnitude of associations between abortion and particular mental health problems across all available studies must then be evaluated.

As a sensitive topic, conducting meaningful, sound research on abortion and mental health is notoriously difficult. Women often never disclose their experiences, are unwilling or unable to revisit their abortions for the purpose of scientific study, and/or fail to continue as participants for any length of time when consent is secured. Perhaps even more serious than this are the well-entrenched, seldom-discussed professional gatekeepers, who encourage agenda-driven research and ignore fundamental scientific principles in order to fill journal articles with “evidence” that abortion poses no risk to women’s health.

Virtually all mainstream professional organizations in psychology and medicine have embraced an unwavering pro-choice political stance on abortion. For example, in 1969 the American Psychological Association (APA) adopted the position that abortion is a civil right, stating: “Whereas, termination of unwanted pregnancies is clearly a mental health and child welfare issue, and a legitimate concern
of APA, be it resolved, that termination of pregnancy be considered a civil right of the pregnant woman, to be handled as other medical and surgical procedures in consultation with her physician."

For decades, the APA has been tightly aligned with major organizations promulgating pro-choice social agendas, such as the American Civil Liberties Union Reproductive Freedom Project, the National Abortion Federation, the National Abortion Rights Action League, the Guttmacher Institute, and Planned Parenthood, among others, frequently composing amicus briefs and offering congressional expert testimony. Martel addressed the APA's position on abortion, noting its promotion of psychological research for the purpose of agenda-driven legislative and broader societal change. Martel further observed that APA's political stance is devoid of strong backing of empirical data. With nearly half a century of abortion advocacy by the most powerful professional psychology organization in the world, incentives to publish confirmatory findings, whatever the scientific costs, are logical, albeit unethical and dangerous to women.

When a one-sided political mission permeates the fabric of major professional organizations across various disciplines, the effects become systemic, influencing far more than the publication decisions of affiliated journals. The training and continuing education of researchers and practitioners, grant-funding priorities, political initiatives, press releases, and ultimately cultural mores are transformed by a process antithetical to science, with a harrowing loss in objectivity.

Many narrative reviews of the empirical literature published in recent years involved specified evaluation criteria that resulted in dismissal of large bodies of evidence with politically unsavory results. To illustrate, the tactics employed by the Royal College of Psychiatrists (RCP) in their 200-plus-page literature review serves as perhaps the highest profile example. This review was pitched as methodologically superior to all previously conducted reviews based on the criteria employed to critique individual studies. However, the quality scales used were not well-validated and required a significant level of subjective interpretation, opening the results to considerable bias. The main problems with the RCP's scales were as follows: 1) the categories used were missing key methodological features, including initial consent-to-participate and retention of participants rates across the study period; 2) the relative importance assigned to the included criteria was arbitrary, as opposed to being based on scientific principles; 3) the specific requirements for assigning a "+" or "-" within the various categories were not provided; 4) the authors failed to explain (similar to their predecessors, Charles et al.) how combinations of pluses and minuses in the distinct categories added up to an overall rating ranging from "Very Poor" to "Very Good."

Incredibly, the 1995 study by Gilchrist et al. received a rating of "good" in the RCP review when very few controls for confounding variables were employed, meaning that the comparison groups likely differed systematically in income, relationship quality (including exposure to domestic violence), social support, and other potentially critical factors. Further, Gilchrist et al. reported retaining only 34.4% of the termination group and only 43.4% of the no-termination group at the end of the study. No standardized measures for mental health diagnoses were employed, and evaluation of the psychological state of patients was reported by general practitioners, not psychiatrists or psychologists. Despite these facts, the study received a mark of "+" for confounder control, a "+" for representativeness and a "+" for validated tools.

Among the scores of studies identified and excluded in the RCP review, the most common reasons were the nebulous "no usable data" and "less [sic] than 90 days follow-up." The latter resulted in elimination of 35 peer-reviewed studies in each of the prevalence, risk factors, and comparison study forms. The RCP authors stated: "Because the review aimed solely to assess mental health problems and substance use, and not transient reactions to a stressful event, negative reactions, and assessments of mental state confined to less [sic] than 90 days following abortion were excluded from the review." p 12

This is highly problematic for various reasons. Elimination of studies only measuring women's mental health for up to 90 days does not consider that many of the women in these studies likely continued to suffer beyond the early assessment. Moreover, when investigating mental health implications of an event, it is logical to measure outcomes soon afterward, as opposed to waiting months or years to gather data. With time, more healing may naturally occur, there may be moderating events, and more confounding variables could be introduced. Finally, focusing only on mental health outcomes that occur later effectively misses the more acute episodes that are treated soon after exposure. Many of the studies that were removed because of brief follow-up had controls for psychological history, and had other strengths; thus, the final set of studies in each section of the RCP review is not representative of the best available evidence.

A large percentage of the eliminated studies coincidentally revealed adverse post-abortion consequences for significant percentages of women. In the category in which the authors sought to derive prevalence estimates, only 34 studies were retained, including 27 without controls for previous mental health. Based on only four studies, authors of the RCP review concluded that abortion is no riskier to women's mental health than unintended pregnancy and delivery.

In 2016 I analyzed the most serious challenges to conducting valid research on the psychology of abortion, explained how each one compromises the quality of evidence, reported on progress in remedying the challenges, and made suggestions for future innovations. I further introduced a rubric that enables one to readily and objectively score an individual study based on the extent to which it successfully addresses the outlined methodological challenges.

After reproducing the scoring rubric revised slightly for clarity (Table 1), a preliminary application to six recently published studies is described (Table 2). The scheme uses nine criteria, and results in individual study scores of zero to 36, with higher scores indicating stronger methodology. The selection includes peer-reviewed journal articles published worldwide between January 2015 and January 2017. Each article examines the association between...
abortion experience, and at least one measure designed to capture adverse mental health outcomes. Three of the studies yielded results suggesting that abortion operates as a significant risk factor for adverse mental health, and three of the studies had contrary results. I am currently undertaking an exhaustive analysis of studies published over the last 25 years applying this scheme. This early snapshot into an approach to providing a rigorous alternative to the subjective, politically motivated reviews populating the literature will also serve to indirectly substantiate views presented above, as studies falling short on essential scientific features are being published in high-profile journals.

Table 1: Scheme for Evaluating Studies on Abortion and Mental Health

<table>
<thead>
<tr>
<th>Criterion</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sample size (S)</td>
<td>50 or fewer</td>
<td>51-199</td>
<td>200-399</td>
<td>400-999</td>
<td>1000 or more</td>
</tr>
<tr>
<td>2. Generalizability (G)</td>
<td>Restricted to one city</td>
<td>2 to 4 cities within 200 miles of each other</td>
<td>5 or more cities in different geographical locations, over 200 miles apart with no evidence the sample represents the population</td>
<td>5 or more cities in different geographical locations, over 200 miles apart with evidence the sample somewhat approximates the population</td>
<td>5 or more cities/ nationally representative/ population based/ large international study including 3 or more nations</td>
</tr>
<tr>
<td>3. Consent to participate rate (CP)</td>
<td>Not available or under 20%</td>
<td>20%-39%</td>
<td>40%-59%</td>
<td>60%</td>
<td>79%</td>
</tr>
<tr>
<td>4. Abortion concealment (C)</td>
<td>Includes women prone to concealment (minors, victims of domestic violence, highly religious or conservative family background)</td>
<td>Concealment rates equivalent to typical studies on abortion</td>
<td>Methodology employed some effort to reduce concealment</td>
<td>Methodology employed extensive strategies to reduce concealment</td>
<td>No concealment/ record-based data/ data secured at an abortion clinic</td>
</tr>
<tr>
<td>5. Attrition rate (longitudinal studies only (A))</td>
<td>High: 44% or less of sample retained</td>
<td>Moderately high: 45-59% of sample retained</td>
<td>Moderate: 60-74% of sample retained</td>
<td>Moderately low: 75-89% of sample retained</td>
<td>Low: 90-100% of sample retained</td>
</tr>
<tr>
<td>6. Control for potentially confounding variables (CC)</td>
<td>No controls for potential confounders</td>
<td>5 or fewer demographic control variables</td>
<td>6 or more controls for several potential confounders not restricted to demographic factors</td>
<td>6 or more controls for several potential confounders, not restricted to demographic factors and including prior mental health</td>
<td>6 or more controls for several potential confounders, not restricted to demographic factors and including prior mental health and pregnancy intendedness</td>
</tr>
<tr>
<td>7. Inclusion of a control group (CG)</td>
<td>No comparison group</td>
<td>Women with no reproductive event or women from the general population</td>
<td>Women who gave birth without intendedness identified</td>
<td>Other form of perinatal loss (miscarriage, stillbirth, adoption placement)</td>
<td>Unintended pregnancy delivered with or without women having actively considered abortion</td>
</tr>
<tr>
<td>8. Strength of measures of mental health outcomes (MMH)</td>
<td>Variables measured with fewer than 10 self-report measures of outcomes with preliminary or no establishment of psychometric properties</td>
<td>Use of self-report measures with 10 or fewer items per variable and some evidence of psychometric properties for the scales used</td>
<td>Use of multiple item self-report measures (10 or more) with extensive well-established psychometric properties</td>
<td>Use of multiple item self-report measures (10 or more) with extensive well-established psychometric properties and another form of data other than self-report</td>
<td>Mental illnesses diagnosed by a trained professional using a well-developed assessment scale or protocol</td>
</tr>
<tr>
<td>9. Prospective data collection (PDC)</td>
<td>One post-abortion assessment</td>
<td>Two or more post-abortion assessments</td>
<td>Two or more assessments, with the first occurring between the time of abortion or within 6 month of the procedure</td>
<td>Pre and post-abortion assessments with one or more post-abortion assessment(s) extending up to a year after the procedure</td>
<td>Pre-abortion assessment(s) and extensive post-abortion assessments, extending from at least a month before to more than a year after the procedure</td>
</tr>
</tbody>
</table>
### Table 2: Application of Evaluation Criteria to Recently Published Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Synopsis</th>
<th>S</th>
<th>G</th>
<th>CP</th>
<th>C</th>
<th>A</th>
<th>CC</th>
<th>CG</th>
<th>MMH</th>
<th>PDC</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biggs et al. [10]</td>
<td>Compared psychological outcomes of women who received abortions just under facility gestational limits with women who wanted an abortion but were denied because they were just over the facility gestational limit. Those denied were subdivided into those who gave birth and those who obtained an abortion subsequently or miscarried. Women denied an abortion, particularly those who later miscarried or had an abortion elsewhere, had the most elevated levels of anxiety and lowest self-esteem and life satisfaction 1 week after being denied an abortion, improving and approaching levels similar to those in the other groups by 6 to 12 months.</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>Gissler et al. [12]</td>
<td>Studied suicide trends after induced abortion in 1987–2012 in Finland. Women with a recent induced abortion still have a 2-fold suicide risk even after current care guidelines include a 2-3 week follow-up session with abortion patients to monitor for mental health disorders.</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>McCarthy et al. [13]</td>
<td>Investigated women with previous pregnancy losses for higher levels of anxiety, depression, stress, and altered behaviors in a subsequent pregnancy. Stress scores were significantly elevated at 15 weeks of gestation in women with both 1 and 2 abortions. Elevated depression was observed in women with one and two prior abortions at 15 and 20 weeks of gestation. Women with two prior abortions displayed increased limiting/resting responses to pregnancy scores at 15 and 20 weeks. No differences were observed in stress or depression scores at 15 weeks between abortion and miscarriage groups; however anxiety was higher in the miscarriage group.</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>26</td>
</tr>
<tr>
<td>Sullins [14]</td>
<td>Examined the links between pregnancy outcomes (birth, abortion, or miscarriage) and mental health outcomes for US women. Abortion was consistently associated with increased risk of mental health disorder. Overall risk was elevated 45%. Risk of mental health disorder with miscarriage was mixed, but also elevated 24% overall. Birth was weakly associated with reduced mental disorders. One-eleventh (8.7%) of the prevalence of mental disorders examined over the period were attributable to abortion.</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>Toffol et al. [15]</td>
<td>Assessed anxiety and quality of life during a one-year follow up after first-trimester abortion. Abortion was associated with an overall reduction of anxiety and improvement in quality of life. High baseline anxiety, history of psychiatric morbidity and smoking among women receiving abortion was persistently associated with high levels of anxiety and poor quality of life.</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Van Ditzhuijzen et al. [16]</td>
<td>Women who had an abortion did not show higher odds for incidence of “any mental disorder,” or mood, anxiety and substance-use disorders, compared to matched controls. Having an abortion was not associated with increased odds for recurrence of the three disorder categories, but for any mental disorder the higher odds in the abortion group remained significant after matching.</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>20</td>
</tr>
</tbody>
</table>

S=Size; G=Generalizability; CP=Consent to Participate; C=Concealment; A=Attrition; CC=Control for Confounding Variables; CG=Control Group; MMH=Measures of Mental Health; PDC=Prospective Data Collection
Evaluation data in Table 2 reveals that the average score for studies with data generally indicating that women who undergo abortions are at increased risk for mental health problems is 28 on a scale of zero to 36; whereas the average score for studies indicating that abortion is not a risk factor for mental health problems is 18.3. These data suggest that a straightforward approach directed toward scrutinizing methodological strengths and weaknesses of individual studies carries considerable potential for bringing clarity to a complex literature, long contaminated by politics. Unlike previous attempts to evaluate key features of individual studies, the scheme employed herein offers a more comprehensive list of methodological elements along with detailed anchoring points, minimizing subjective interpretation.

Future efforts to apply the scale to hundreds of studies from the world literature should lead to discerning the relative importance of each scale component with rankings and scoring weights deemed scientifically justified. Such work will likely result in articulation of study features that must receive a minimum score in order for the study to be taken seriously despite strengths in other areas. For example, the recently published JAMA Psychiatry article by Biggs et al., described in Table 2, involved shockingly simplistic measures, with two variables containing only six items (depression and anxiety) and two variables (satisfaction with life and self-esteem) measured with single items. Such an approach is indefensible given that there are numerous psychometrically sound multiple-item inventories and clinical interview protocols available in the literature. This invites the questions: what were the researchers measuring exactly? How did their data relate to the potentially profound adverse reactions that women experience? Does it matter how many cities were sampled or how many assessment points were used if what is being assessed is relatively meaningless?

Future refinement of the scale proposed herein may necessitate identification of central study features and/or a baseline total cut score for studies to be considered worthy contributions to literature reviews, policy, and practice. If an objective rubric approach gains wide-scale acceptance over time, journal editors and manuscript reviewers will ideally use a similar approach, reducing the percentage of flawed studies in the professional literature.

Most studies published today include some controls for third variables. However, a more sophisticated analysis of the many factors differentiating women who choose abortion from those who do not is needed in order to achieve consensus on the best set of third variables to use in future studies. Researchers studying this typically attempt to include assessments of age, ethnicity, pregnancy timing, reproductive history, relationship history/marital status, socio-economic status, partner support, pregnancy intention, and prior mental health when possible, as these are widely recognized predictors of the choice to have an abortion, and have been associated with differences in post-abortion psychological function. Other variables that have demonstrated discriminatory power include satisfaction with pre-abortion counseling, exposure to domestic violence as a child and as an adult, religious affiliation, and liberal vs. conservative political beliefs. Personal variables that have consistently surfaced in the literature include beliefs about the humanity of the fetus, bonding to the fetus, decision ambivalence, and pressure to abort. Future studies continuing the tradition of presenting statistically adjusted and non-adjusted results using the above variables and others will help evaluate work that has included, or failed to include, the most essential covariates.

Identification of the best measures of mental health to use in future studies is another area needing significant development and refinement. Most studies contain self-report instruments with reasonable psychometric evidence documenting scale reliability and validity. However, ideally multi-modal assessment forms are used with data from mental health professionals, partners, and family members to arrive at a deeper understanding of how abortion potentially affects women and those they love. The need for a well-developed measure of a broad range of abortion’s potential adverse outcomes has received only scant professional attention to date. The literature is in need of a comprehensive scale for addressing abortion effects, and it should include less commonly assessed adverse mental, behavioral, and relational variables, such as anger, detachment and loneliness, self-doubt, negative self-appraisals, risky behaviors, and difficulty relaxing/staying busy to avoid thinking about the abortion, among others.

Most outcomes we see in the literature pertain to mood, anxiety, and substance abuse disorders, along with suicidal thinking and behaviors. However, many women have adverse effects without a full-blown mental illness. They have only a few symptoms of a particular disorder, yet the impact is real, disturbing, and potentially debilitating. We need to understand what the most commonly experienced symptoms of various disorders are following abortion, and the extent to which they reduce the quality of women’s lives. For example, an individual may report excessive worry and anxiety characteristic of Generalized Anxiety Disorder, but she does not meet criteria for the disorder without impairment in social, occupational, or other areas of functioning. This woman may be engaged in relationships and productive work, but this does not mean she is not suffering, or that she is living a full life.

Many women who decide to abort do so without a thorough understanding of the potential emotional ramifications of the procedure, despite widespread desire to be informed. In 2006 colleagues and I published a study in the Journal of Medical Ethics, and our data revealed 95 percent of a socio-demographically diverse group of women wanted to be informed of all possible complications associated with drugs, surgery, and/or other elective treatments, including abortion.
In other areas of medicine, best-practice protocols evolve within the various specialties among researchers, editors, and leaders of professional organizations, working in concert toward the goal of quality care. Local physicians and allied health professionals largely trust the professional organizations to package and disseminate the latest knowledge to improve service provision. In sharp contrast, the only way to guarantee women will be provided evidence-based abortion services is for states to pass laws. Legislators nationwide have responded to women’s needs for information about the abortion decision’s gravity by enacting laws mandating provision of accurate information regarding the risks involved, substantive pre-abortion counseling, and time to reflect on the decision prior to consent.

Fortunately, 22 states currently require pre-abortion counseling on possible psychological responses to the procedure, nine of which emphasize negative emotional reactions. This process is often extremely time-consuming and costly, with experts on both sides debating the scientific merits of peer-reviewed studies containing dramatically different results and conclusions. As professionals, we have the ethical responsibility to conduct sophisticated studies and perform trustworthy reviews of the peer-reviewed literature in order to simplify and support legislative efforts. Only then will the process of proposing and passing legislation that accurately reflects and respects the fullness of women’s emotional responses to unplanned pregnancy and abortion be expeditious and truly evidence-based.

Priscilla K. Coleman, Ph.D., is Professor of Human Development and Family studies, 1128 Eppler North, Bowling Green State University, Bowling Green, OH 43403. Contact: pcolembgsu.edu.

REFERENCES


