

SHORT COMMUNICATION

TITLE: Long-term Incidence and Recurrence of Common Mental Disorders after Abortion. A Dutch Prospective Cohort Study.

AUTHORS

Jenneke van Ditzhuijzen, PhD ^{a*}

Margreet ten Have, PhD ^b

Ron de Graaf, PhD ^b

Carolus H.C.J. van Nijnatten, PhD ^a

Wilma A.M. Vollebergh, PhD ^a

^a Department of Interdisciplinary Social Science, Utrecht University, Utrecht, the Netherlands

^b Netherlands Institute of Mental Health and Addiction (Trimbos Institute), Utrecht, the Netherlands

* Corresponding author. Telephone: +31 30 2534742, E-mail: j.m.vanditzhuijzen@uu.nl, Postal address: P.O. Box 80.140, 3508TC, Utrecht, the Netherlands.

ABSTRACT

In a previous study (Van Ditzhuijzen et al., 2017) we investigated the incidence and recurrence of mental disorders 2,5 to 3 years post-abortion. The aim of the current study was to extend these findings with longer term follow up data, up until 5 to 6 years post-abortion. We compared data of women who had had an abortion of the Dutch Abortion and Mental Health Study (DAMHS) to women who did not have an abortion from the Netherlands Mental Health Survey and Incidence Study-2 (NEMESIS-2) ($N_{\text{total}} = 2\,227$). We used 1-to-1 matching on background confounding variables and measured post-abortion incidence and recurrence of common DSM-IV mental disorders (mood, anxiety, and substance use disorders) using the Composite International Diagnostic Interview (CIDI) version 3.0. After matching on confounding variables, abortion did not increase the likelihood that women had incident or recurrent mental disorders in the 5 to 6 years post-abortion (any incident mental disorder: $OR=3.66$, $p=.16$; any recurrent mental disorder: $OR=.22$, $p=.47$). We found no evidence that experiencing an abortion increases the risk on new or recurrent mental disorders on the longer term.

Key words: induced abortion; unwanted pregnancy; incidence; recurrence; common mental disorders; coarsened exact matching.

Credible review studies have stated that there is no evidence that abortion causes adverse mental health outcomes (APA, 2008, Charles et al., 2008; NCCMH, 2011). However, the field is characterized by methodological limitations such as insufficient confounder control, and conclusions have mostly been based on cross-sectional or short-term follow-up studies. It could therefore be argued that mental health problems that might arise only years after an abortion, are being missed in research; but the few high-quality studies that have focused on long-term postabortion mental health have found no evidence for this (e.g., Biggs et al., 2015; 2017; Munk-Olsen et al., 2011). However, so far no studies have measured long-term incidence and recurrence of a wide array of clinical mental disorders. The current study aims to fill this gap.

In an earlier study, we found that abortion was not associated with higher incidence of disorders in the 2.5 to 3 years post-abortion, but women who had an abortion seemed at slightly increased risk for recurrence of any mental disorder (Van Ditzhuijzen et al., 2017). In the current study, we investigated whether these results hold in the long term. The aim was to answer the question whether the life event of terminating an unwanted pregnancy increases the risk on the development of clinical DSM-IV mental disorders 5 to 6 years postabortion.

METHODS

Study design

In this prospective three-wave cohort study we compared women who had an abortion with women from the general population who never had an abortion. In research like this, it is impossible to use random assignment to treatment (abortion) or control condition, which implies that selectivity and confounding are inherent to this type of research. We used Coarsened Exact Matching (CEM) to deal with this general problem (Stuart, 2010; Rosenbaum and Rubin, 1983; Dehejia and Wahba, 2002; Cook et al., 2008), hereby inducing balance in the distributions of the covariates. No additional adjustment for covariates was needed, besides for timing between measurement waves. We demonstrated the effect of matching with CEM on the distribution of covariates in an earlier study (Van Ditzhuijzen et al., 2017).

Setting

The Dutch Abortion and Mental Health Study (DAMHS) participants were recruited from seven Dutch abortion clinics spread out over the country. Shortly after the abortion procedure, abortion clinic staff members asked the women to read the research flyer, complete a reply card including a consent-to-contact statement or a non-response form, and deposit this card in a locked mailbox.

Participants were interviewed by professionally trained female interviewers at three time points. First, approximately 20 to 40 days post-abortion, between April 2010 and January 2011 (T0), then between December 2012 and November 2013, which was on average 2.7 years later (T1), and lastly, between April 2015 and November 2015 (T2), on average 4.9 years postabortion. The mean duration of the interviews was around 2.5 h (T0) and 1.5h (T1 and T2). Participants received a gift card of 50 Euros for each interview.

The reference cohort was taken from the Netherlands Mental Health Survey and Incidence Study -2 (NEMESIS-2). NEMESIS-2 Participants were first interviewed (T0) between November 2007 and July 2009. First follow-up (T1) was 3 years later between October 2010 and May 2012, and second follow-up 6 years later between November 2013 and June 2015 (T2). The selection strategy of NEMESIS-2 is described extensively elsewhere (De Graaf et al., 2010).

Both studies were approved by a local medical ethics committee. In both studies, written informed consent was obtained before each interview, and all interviews were fully laptop-assisted, face-to-face, and done by professionally trained interviewers.

Participants

The DAMHS study enrolled Dutch-speaking women of 18 to 46 years old, obtaining an abortion (medical or aspiration, until a maximum of 22 weeks) for an unwanted pregnancy, without clear fetal or maternal medical indications. We included from NEMESIS-2 women in the same age range as in DAMHS (18-46), who reported they never experienced abortion. Participants were updated through e-mail or regular mail about the study planning, and re-approached to schedule the follow-up interviews by telephone, e-mail, and text messages.

Measures

Mental disorders

In both cohorts, presence of lifetime and last-year DSM-IV disorders was assessed at T0 with the Composite International Diagnostic Interview (CIDI) version 3.0. (Alonso et al., 2004; Haro et al., 2006). At T1 and T2, presence of mental disorders since the last interview was assessed. The following disorders were included: mood disorders (major depression, dysthymia, bipolar disorder); anxiety disorders (panic disorder, agoraphobia, social phobia, specific phobia, generalized anxiety disorder); and substance use disorders (alcohol/drug abuse and dependence). For incidence, women who never had one or more disorders within a specific disorder category before T0 were included in the at risk group. For recurrence, women who had one or more disorders within a specific disorder category in their lifetime, but not in the 12 months before T0, were included. By excluding 12-month prevalent cases, women who had a disorder around the time of the abortion were excluded, this way the correct order of the events (abortion and mental disorder) was ensured. In addition to the three disorder categories, we also included the aggregate measure any mental disorder. For this variable, the at risk group for incidence consisted of women who never had a disorder in any of the three disorder categories. Likewise, recurrence referred to disorders in any of the three categories.

Covariates

Demographic variables assessed at baseline were age category (18-24, 25-34 and 35-46 years); living situation (with or without a partner); having children (yes or no); western or non-western ethnicity based on the definition of Statistics Netherlands (Stronks et al., 2009); considers herself religious (yes or no); employment situation (paid job or not); education level; and urbanicity of place of residence (urban or rural). Furthermore, we also included childhood abuse as a covariate, because it has been found that childhood abuse can predispose women for both mental disorders (Kessler et al., 1997; Green et al., 2010), and abortion (Boden et al., 2009; Steinberg and Tschann, 2013). Following other NEMESIS-2 studies (De Graaf et al., 2010), childhood abuse was scored '1' when one or more types

of abuse had happened at least once (sexual abuse) or more than once (emotional, psychological, and physical abuse) before the age of 16.

Statistical analysis

First, multivariable logistic regression analyses were performed with unmatched data, only adjusting for the difference in time interval between T0 and T2. Based on the literature, we then identified potential covariates associated with both abortion and mental health. For each of these variables, we individually tested whether they were associated with the predictor (abortion versus control cohort), and second, whether they were associated with the various outcome variables (incidence or recurrence of the four categories of mental disorders at T2). We selected only variables that were associated with the predictor and at least one outcome measure. All variables selected for matching (age category, living situation, having children, western or non-western ethnicity, employment situation, living in an urban environment, and a history of childhood abuse) were categorical. Matching was implemented in the Coarsened Exact Matching package (Iacus et al., 2011, 2012) for SPSS. We matched participants 1-to-1, and included only exact matches (282 pairs). This implies that pairs of women from both cohorts are formed that are identical on all selected covariates, and differ only with respect to having had an abortion. Matching was done on T0 data, before attrition. After this, multivariable logistic regression analyses were performed for each disorder on T2 (both incidence and recurrence), for matched data; adjusting for the length of the time interval between T0 and T2 in years. Loss to follow-up was addressed in an attrition analysis. Testing was two-sided, and statistical significance was considered to be $p < .05$. Analyses were performed using SPSS version 22.

RESULTS

Participants

At T0, we included 325 participants in DAMHS. The initial response rate was 36.5%, mainly because we were unable to schedule interviews with women who had agreed to participate. Participant flow and response analysis results at T0 are described extensively in an earlier study (Van Ditzhuijzen et

al., 2013). Of the initial 325 participants in DAMHS, 264 (81.2%) were re-interviewed at T1, 40 women could not be traced, 13 refused participation, and 8 women did not show up at the interview. At T2, 231 out of 264 (87.5%) women were re-interviewed, 15 could not be traced, 9 refused participation, and 9 did not show up at the interview.

NEMESIS-2 had an initial response rate at T0 of 65.1%. We included 1902 women in our T0 reference cohort from NEMESIS-2. Of the 1902 women selected from NEMESIS-2 for this study at T0, 1496 were re-interviewed at T1 (78.7%), and 1297 at T2 (87.2%). Of these, 14 had undergone abortion between T0 and T2. They were excluded, leaving 1283 participants in the reference group at T2.

Demographic characteristics of the DAMHS and the NEMESIS-2 cohort at T0, as well as a response analysis, are described in Van Ditzhuijzen et al., 2013. Of the 325 cases at baseline in DAMHS (exposed), 282 could be matched to an exact NEMESIS-2 match (unexposed) on selected background variables. After attrition, the DAMHS cohort consisted of 210 cases and NEMESIS-2 of 179 cases.

Attrition

Overall, attrition between T0 and T2 was comparable for both cohorts. Attrition analysis showed that in DAMHS, women aged 18-24 at baseline dropped out more often than women aged 35-46 at baseline (OR=4.67, CI=1.88-11.65, $p=.001$), as well as women of non-western origin compared to women of Dutch or other western ethnicity (OR=2.15, CI=1.08-4.27, $p=.03$). In NEMESIS-2, women without children at T0 (OR=1.55, CI=1.20-2.00, $p=.001$), women of non-western ethnicity (OR=2.22; CI=1.50-3.29, $p<.000$), and women with the two lowest levels of education (OR=1.97, CI=1.06-3.64, $p=.03$; and OR=1.60, CI=1.21-2.11, $p=.001$) discontinued more often. We also analyzed whether the mental disorder categories studied here were associated with attrition (while controlling for the demographic variables that were associated with attrition in one or both cohorts). In NEMESIS-2, but not in DAMHS, women who had anxiety disorders at T0 discontinued less often than women who did not have anxiety disorders (OR=.73, CI=.53-.99, $p=.04$). Mood disorders, substance abuse disorders, and the aggregate measure any mental disorder were not associated with attrition in the two cohorts.

Incidence

Before matching, participants who had an abortion had higher odds for incidence of the aggregate measure any mental disorder (OR=3.88, CI=1.34-11.19, $p=.01$), and for substance use disorders (OR=5.02, CI=1.14-22.23, $p=.03$) compared to participants in the no-abortion group. The odds for incidence of mood and anxiety disorders were not higher for the abortion group compared to controls before matching (see Table). After matching, the differences in incidence between the cohorts decreased substantially to non-significance. However, for mood disorders the difference between the two cohorts remained close to significant, but the confidence interval became much wider, therefore we believe this might be related to outliers in the DAMHS cohort.

Recurrence

Before matching, participants in the DAMHS cohort had slightly higher odds for substance use disorders only (OR=17.20, CI=1.01-292.24, $p=.049$) than participants in NEMESIS-2, but this was not the case for mood disorders, anxiety disorders, or any mental disorder (see Table 1). Because there were no recurrent cases for substance use disorders left after matching and attrition in the NEMESIS-2 cohort, we could not calculate the B-parameter and fixed it at 15.

DISCUSSION

In the unmatched (unadjusted) data, we found differences between cohorts for long-term incidence and recurrence of substance use disorders, and in incidence of any mental disorder, but not for other categories of disorders. However, these differences were strongly reduced after 1-to-1 matching on confounding variables. The initial differences seem not attributable to the abortion itself, but rather related to confounding variables.

Comparing the current findings to the findings after 2.5 to 3 years post-abortion (Van Ditzhuijzen et al., 2017), it is noteworthy that even the 'raw', unmatched differences between the cohorts became smaller over time, especially for incidence. This is in line with findings that symptoms of mental

health disorders decrease over time after an abortion (Foster et al., 2015). At both time points in our study, matching on confounders lead to disappearance of the effects for incidence; with the exception of incident mood disorders, which remained close to significant after matching. The large confidence interval of mood disorders shows us that we should not interpret this almost-significant result. For recurrence, a small effect was visible in the aggregate measure any mental disorder on the shorter term, but this dissipated after 5 to 6 years. All in all, these findings show that it is unlikely that abortion is causing common mental disorders, which is in line with conclusions of various review studies (APA, 2008; Charles et al., 2008, NCCMH, 2011).

The following limitations should be mentioned. Because of the relatively low numbers in this study, also due to the matching procedure and loss to attrition, confidence intervals of the odds ratios were relatively wide, especially for recurrence. We were also not able to calculate point estimates for recurrent substance use disorders after matching. However, the absence of strong abortion-related effects is not caused by the matching procedure; the unmatched data already revealed few differences. Second, we cannot be certain that we matched on all factors that were associated with abortion and subsequent mental disorders. Even though matching has the potential to reduce unmeasured confounding (Stürmer et al., 2010), it is unlikely that it removed all confounding completely. Third, we only included adult (18+) women, therefore these results do not necessarily generalize to adolescents. However, other research suggests that there is no reason to expect different results for younger women (Warren et al., 2010; Stotland, 2011; Gomez, 2017). Fourth, just like the experimental design, the matching design has the advantage that internal validity is strong, but external validity is limited. Also, abortion studies usually have low participation rates (APA, 2008; Foster et al., 2015), and the current study is no exception. Even though the matching dealt with selectivity, we should remain prudent in generalizing the results.

To our knowledge, we have been the first to use 1-to-1 exact matching in examining the association between abortion and mental disorders. We confirmed that pre-existing differences are crucial, as they strongly confound the associations between abortion and mental health. If confounders are not

controlled for in a stringent manner, differences between groups are likely to be falsely attributed to the abortion (APA, 2008; Charles et al., 2008; Van Ditzhuijzen et al., 2013; 2017).

In the Netherlands, abortion is free, legal, and available up until 22 weeks of gestation; but at the same time the abortion rate is among the lowest in the world. Even though women with a psychiatric history seem to be overrepresented in Dutch abortion care, the current study also confirms that abortion does not, in itself, increase the risk on mental disorders.

Table 1. Five to Six Year Incidence and Recurrence of Mental Disorders by Study Group, Unmatched and Matched.

	Total No. (%)	DAMHS No. (%)	NEMESIS-2 No. (%)	OR (95% CI)	P
Incidence, non-matched					
Mood disorders	127 / 1106 (11.5)	28 / 137 (20.4)	99 / 969 (10.2)	2.94 (.97-8.87)	.06
Anxiety disorders	130 / 1123 (11.6)	27 / 137 (19.7)	103 / 986 (10.4)	1.97 (.63-6.12)	.24
Substance use disorders	58 / 1343 (4.3)	16 / 180 (8.9)	42 / 1163 (3.6)	5.02 (1.14-22.23)	.03
Any common mental disorder ^a	153 / 850 (18.0)	29 / 77 (37.7)	124 / 773 (16.0)	3.88 (1.34-11.19)	.01
Incidence, matched (282^b pairs)					
Mood disorders	42 / 254 (16.5)	24 / 121 (19.8)	18 / 133 (13.5)	6.53 (.99-43.19)	.052
Anxiety disorders	42 / 253 (16.6)	23 / 124 (18.5)	19 / 129 (14.7)	1.14 (.14-9.13)	.90
Substance use disorders	22 / 327 (6.7)	15 / 164 (9.1)	7 / 163 (4.3)	7.35 (.52-104.98)	.14
Any common mental disorder	45 / 171 (26.3)	25 / 68 (36.8)	20 / 103 (19.4)	3.66 (.54-24.65)	.18
Recurrence, non-matched					
Mood disorders	89 / 413 (21.5)	24 / 79 (30.4)	65 / 334 (19.5)	.66 (.09-4.55)	.67
Anxiety disorders	44 / 246 (17.9)	15 / 52 (28.8)	29 / 194 (14.9)	6.16 (.93-40.76)	.06
Substance use disorders	27 / 169 (16.0)	15 / 45 (33.3)	12 / 124 (9.7)	17.20 (1.01-292.24)	.049
Any common mental disorder	142 / 504 (28.2)	39 / 88 (44.3)	103 / 416 (24.8)	1.07 (.26-4.47)	.93
Recurrence, matched (282^b pairs)					
Mood disorders	29 / 86 (33.7)	23 / 59 (39.0)	6 / 27 (22.2)	.25 (.00-42.98)	.60
Anxiety disorders	21 / 59 (35.6)	13 / 40 (32.5)	8 / 19 (42.1)	.03 (.00-6.22)	.20
Substance use disorders	13 / 37 (35.1)	13 / 29 (44.8)	0 / 8 (0)	Fixed ^c	
Any common mental disorder ^a	49 / 99 (49.5)	37 / 68 (54.4)	12 / 31 (38.7)	.22 (.00-13.15)	.47

^a The number of cases for any common mental disorder is not equal to the sum of the cases of the disorder categories, because participants were considered recurrent if they had a lifetime history of any disorder before T0 (but not in the last year before T0) and a disorder between T0 and T2, regardless of disorder category (for instance, participants with a lifetime mood disorder could develop an anxiety disorder after T0, which was considered recurrent in terms of any common mental disorder).

^b Matching was done with T0 data, before attrition; this is why the number of pairs after matching is larger than the smallest sample (DAMHS) at T2.

^c As there were no recurrent cases for substance use disorders in the NEMESIS-2 cohort, this B-parameter was fixed at 15.

Note. At risk groups for *Incidence* consist of women who never had the disorder before T0 and were interviewed at both waves; *incident* cases are those participants of the at risk group who developed one or more disorders within the disorder category between T0 and T2. At risk groups for *Recurrence* consist of women who had one or more disorders in their lifetime, but not in the last year before T0, and were interviewed at all of the three waves; *recurrent* cases are those participants of the at risk group who developed one or more disorders within the disorder category between T0 and T2. Odds ratios are adjusted for variation in the time interval between T0 and T2. Variables for matching were categorical and all associated with the predictor and at least one of the outcome variables of either incidence or recurrence: age category, living situation, western versus non-western ethnicity, children, urbanicity of residence, employment situation, and history of childhood abuse.

CONTRIBUTORS

Authors Vollebergh and Van Nijnatten designed the study, in cooperation with authors De Graaf and Ten Have. Author Van Ditzhuijzen managed the literature searches, data collection and analysis, supervised by author Ten Have. Author Van Ditzhuijzen wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

ACKNOWLEDGEMENTS

We thank Ellen Wittkamp, MSc, who assisted in the coordination of the T2 data collection; Peter Lugtig, PhD, for statistical advice; and all the interviewers and participants, without whom this study could not have been conducted.

FINANCIAL SUPPORT

This work was supported by the Dutch Ministry of Public health, Welfare and Sports, via the Netherlands Organization for Health Research and Development (ZonMw) [grant number 127000003].

REFERENCES

- Alonso, J., Angermeyer, M.C., Bernert, S., Bruffaerts, R., Brugha, T.S., Bryson, H., et al. Vollebergh W.A.M., 2004. Sampling and methods of the European study of the epidemiology of mental disorders (ESEMeD) project. *Acta Psychiat. Scand. Suppl.*; 420: 8-20.
- American Psychological Association (APA), Task Force on Mental Health and Abortion, 2008. Report of the APA task force on mental health and abortion. Author, Washington, DC. <http://www.apa.org/pi/wpo/mental-health-abortion-report.pdf> (accessed 18 July 2017).
- Biggs, M.A., Neuhaus, J.M., Foster, D.G., 2015. Mental health diagnoses 3 years after receiving or being denied an abortion in the United States. *Am. J. Public Health.* 105, 2557-2563.
- Biggs, M.A., Upadhyay, U.D., McCulloch, C.E., Foster, D.G., 2017. Women's mental health and well-being 5 years after receiving or being denied an abortion. A prospective, longitudinal cohort study. *JAMA Psychiatry.* 74, 169-178.
- Boden, J.M., Fergusson, D.M., Horwood, L.J., 2009. Experience of sexual abuse in childhood and abortion in adolescence and early adulthood. *Child Abuse Neglect.* 33, 870-876.
- Charles, V.E., Polis, C.B., Sridhara, S.K., Blum, R.W., 2008. Abortion and long-term mental health outcomes: A systematic review of the evidence. *Contraception.* 78, 436-450.
- Cook, T.D., Shadish, W.R., Wong, V.C., 2008. Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *J. Policy Anal. Manag.* 27, 724-750.
- De Graaf, R., Ten Have, M., Van Dorsselaer, S., 2010. The Netherlands Mental Health Survey and Incidence Study-2 (NEMESIS-2): Design and methods. *Int. J. Meth. Psychiat. Res.* 19, 125-141.
- Dehejia, R.H., Wahba, S., 2002. Propensity score-matching methods for nonexperimental causal studies. *Rev. Econ. Stat.* 84, 151-161.

- Foster, D.G., Steinberg, J.R., Roberts, S.C., Neuhaus, J., Biggs, M.A., 2015. A comparison of depression and anxiety symptom trajectories between women who had an abortion and women denied one. *Psychol. Med.* 45, 2073-2082.
- Gomez, A.M., 2017. Abortion and subsequent depressive symptoms: An analysis of the National Longitudinal Study of Adolescent Health. *Psychol. Med.* [epub ahead of print], 1-11.
- Green, J.G., McLaughlin, K.A., Berglund, P.A., Gruber, M.J., Sampson, N.A., Zaslavsky, A.M., Kessler, R.C., 2010. Childhood adversities and adult psychiatric disorders in the National Comorbidity Survey Replication (NCS-R) I: Associations with first onset of DSM-IV Disorders. *Arch. Gen. Psychiat.* 67, 113-123.
- Haro, J.M., Arbabzadeh-Bouchez, S., Brugha, T.S., de Girolamo, G., Guyer, M.E., Jin, R., et al. Kessler, R.C., 2006. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health Surveys. *Int. J. Meth. Psychiat. Res.* 15, 167-180.
- Iacus, S.M., King, G., Porro, G., 2011. Multivariate matching methods that are monotonic imbalance bounding. *J. Am. Stat. Assoc.* 106, 345-361.
- Iacus, S.M., King, G., Porro, G., 2012. Causal inference without balance checking: Coarsened Exact Matching. *Polit. Anal.* 20, 1-24.
- Kessler, R.C., Davis, C.G., Kendler, K.S., 1997. Childhood adversity and adult psychiatric disorder in the US National Comorbidity Survey. *Psychol. Med.* 27, 1101-1119.
- Munk-Olsen, T., Laursen, T., Pedersen, C.B., Lidegaard, Ø., Mortensen, P., 2011. Induced first-trimester abortion and risk of mental disorder. *N. Engl. J. Med.* 364, 332–339.
- National Collaborating Centre for Mental Health (NCCMH), 2011. Induced abortion and mental health: A systematic review of the mental health outcomes of induced abortion, including their prevalence and associated factors. Academy of Medical Royal Colleges, London.

http://www.aomrc.org.uk/doc_view/9432-induced-abortion-and-mental-health (accessed 18 July 2017).

Rosenbaum, P., Rubin, D., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 70, 41-55.

Steinberg, J.R., Tschann, J.M., 2013. Childhood adversities and subsequent risk of one or multiple abortions. *Soc. Sci. Med.* 81, 53-59.

Stotland, N.L., 2011. Induced abortion and adolescent mental health. *Curr. Opin. Obstet. Gynecol.* 23, 340-343.

Stronks, K., Kulu-Glasgow, I., Agyemang, C., 2009. The utility of 'country of birth' for the classification of ethnic groups in health research: The Dutch experience. *Ethn. Health*. 14, 255-269.

Stuart, E.A., 2010. Matching methods for causal inference: A review and look forward. *Stat.Sci.* 25,1-21.

Stürmer, T., Rothman, K.J., Avorn, J., Glynn, R.J., 2010. Treatment effects in the presence of unmeasured confounding: Dealing with observations in the tails of the propensity score distribution – A simulation study. *Am. J. Epidemiol.* 172, 843-854.

Van Ditzhuijzen, J., Ten Have, M., De Graaf, R., Lugtig, P., Van Nijnatten, C.H.C.J., Vollebergh, W.A.M., 2017. Incidence and recurrence of common mental disorders after abortion: Results from a prospective cohort study. *J. Psychiat. Res.* 84, 200-206.

Van Ditzhuijzen, J., Ten Have, M., De Graaf, R., Van Nijnatten, C.H.C.J., Vollebergh, W.A.M., 2013. Psychiatric history of women who have had an abortion. *J. Psychiat. Res.* 47, 1737-1743.

Warren, J.T., Harvey, S.M., Henderson, J.T., 2010. Do depression and low self-esteem follow abortion among adolescents? Evidence from a national study. *Persp. Sex. Reprod. Health.* 42, 230-235.