Gynecologic Oncology 152 (2019) 533–539

Contents lists available at ScienceDirect

Gynecologic Oncology

journal homepage: www.elsevier.com/locate/ygyno

Review Article

Measuring cause-and-effect relationships without randomized clinical trials: Quasi-experimental methods for gynecologic oncology research

Haley A. Moss [a,*], Alexander Melamed [b], Jason D. Wright [c]

[a] Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, NC, United States of America
[b] Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, Massachusetts General Hospital, Boston, MA, United States of America
[c] Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, Columbia University College of Physicians and Surgeons, New York, NY, United States of America

HIGHLIGHTS

- Clinical research in gynecologic oncology has seen a proliferation of observational studies.
- Observational studies comparing the effectiveness of treatments are susceptible to bias from unmeasured confounding.
- Quasi-experimental methodology can estimate unbiased treatment effects when confounders are not measured.

ARTICLE INFO

Article history:
Received 13 August 2018
Received in revised form 26 October 2018
Accepted 5 November 2018

Keywords:
Research methods
Quasi-experimental research design
Gynecologic oncology

ABSTRACT

Clinical research in gynecologic oncology has seen a proliferation of studies that investigate the effectiveness of treatments using existing data sources such as cancer registries, electronic health records, and insurance claims. These observational studies are often feasible when randomized trial may not be, and may be more generalizable than randomized trials, because of greater diversity in the study populations. While statistical methods such as multivariable regression, matching, stratification, and weighting can adjust for the confounding in observational studies, statistical adjustment cannot control for confounders that are unmeasured in the data. Observational studies comparing the effectiveness of treatments for gynecologic malignancies are susceptible to bias from unmeasured confounding because factors like functional status, frailty and disease burden, which influence treatment selection and outcome, are often not reported in existing data sources. Like randomized trials, quasi-experimental designs attempt to account for both measured and unmeasured confounding by exploiting natural experiments arising in the real world. These methods are underutilized in gynecologic oncology research and are particularly relevant to studies that use large datasets to study the effectiveness of treatments. In this review, we consider methodological challenges that arise in the analysis of non-randomized studies, and describe how application of quasi-experimental methodology can estimate unbiased treatment effects even in the presence of unmeasured confounders.

Contents

1. Introduction . . . 534
2. Methodological challenges of non-random treatment assignment . . . 534
3. Examples of quasi-experimental methods . . . 535
   3.1. Regression discontinuity design . . . 535
   3.2. Difference-in-differences . . . 536
   3.3. Interrupted time series . . . 536
   3.4. Instrumental variables . . . 536
4. Selection and implementation of methods . . . 538
5. Conclusion . . . 538
Author contribution . . . 538

* Corresponding author at: Duke University Medical Center, Division of Gynecologic Oncology, Box 3079, Durham, NC 27710, United States of America.
E-mail address: haley.moss@duke.edu (H.A. Moss).

https://doi.org/10.1016/j.ygyno.2018.11.006
0090-8258/

Conflict of interest . . . 538
References. . . . 538

1. Introduction

Quai-experiments are studies that seek to answer causal hypothesis without randomization. While statistical techniques such as multivariable regression, stratification, matching, and weighting can elucidate causal effect in observational data when all confounders are observed [1], in practice unmeasured confounding is a major obstacle to real-world comparisons of the effectiveness of cancer treatments [2]. In observational studies, treatment decisions may be affected by patient characteristics such as age, medical co-morbidities, performance status and disease characteristics. Factors influencing treatment decisions may be closely associated with the outcome of interest. For example, many observational studies have suggested that survival for neoadjuvant chemotherapy is inferior to primary surgery. However, clinician perception of a poor prognosis may influence the provider's decision to proceed with upfront chemotherapy. Therefore, prognosis is a potential confounder in the relationship between treatment modality and patient survival. (Fig. 1) [3] Quasi-experimental study designs such as instrumental variables analysis, regression discontinuity designs, difference-in-differences, and interrupted time series, can estimate unbiased treatment effects from observational data even when important confounders are not measured [4]. These methods are particularly relevant for studies comparing the effectiveness of treatments in existing data sources such as cancer registries, electronic health records, and insurance claims data. This article describes the challenges of non-randomized treatment assignment to investigators seeking to evaluate the effects of interventions in observational data, and outlines quasi-experimental study designs which may help to overcome these obstacles.

2. Methodological challenges of non-random treatment assignment

The emergence of comparative effectiveness research (CER) has prompted renewed interest in research methodology. While randomized controlled trials (RCTs) are considered the "gold standard" for assessing the causal effect of an intervention, RCTs have a number of important limitations. First, RCTs typically have restrictive eligibility criteria and include highly selected patients often from a limited number of specialized centers who may not resemble the population in routine clinical practice. To be useful, the results of an RCT must be generalizable to "real world" populations. Second, RCTs are often not practical. Expense, ethical considerations, lack of a sufficient number of patients, and lack of patient acceptance are common barriers to the conduct of RCTs [2]. Clinical scenarios in which an RCT is not possible are common in gynecologic oncology.

To address the limitations of RCTs, CER places a strong focus on alternative methods of gathering evidence, particularly observational studies. An observational study is an investigation that compares outcomes among groups of patients whose treatments are determined by some process other than randomization by an investigator [2]. Treatment assignment is in part determined by natural variations in clinical practice. Investigators can leverage this variation in care to compare outcomes among patients who received different treatments.

Unlike RCTs which typically include a highly selected group of patients, observational studies often include a large number of diverse patients, which makes them more generalizable to real world populations. Retrospective observational studies usually rely on existing data and can thus be completed more quickly and at lower cost than RCTs. Observational studies require that the treatment of interest is already utilized in clinical practice, thus these studies are not feasible for drugs, devices or procedures not already in use.

The most important limitation to the validity of observational studies is selection bias. Selection bias occurs when non-random treatment allocation results in differences between the comparison groups of an observation study. Selection bias commonly results when characteristics associated with underlying health status or disease severity affect the selection of a given treatment, so called confounding by indication [2]. When a characteristic is associated with both the treatment (independent variable) and the outcome (dependent variable) it is called a confounder. While some confounders are readily identifiable and are measured in a given dataset, important confounders may be unmeasured.

Because of randomization, RCTs are free of selection bias and all measured and unmeasured confounders should be balanced between the control and experimental groups. As the two groups differ only in the assigned treatment, RCTs can estimate the causal effect of an intervention on outcomes. In contrast, observational studies can document associations between an exposure and an outcome, but in the presence

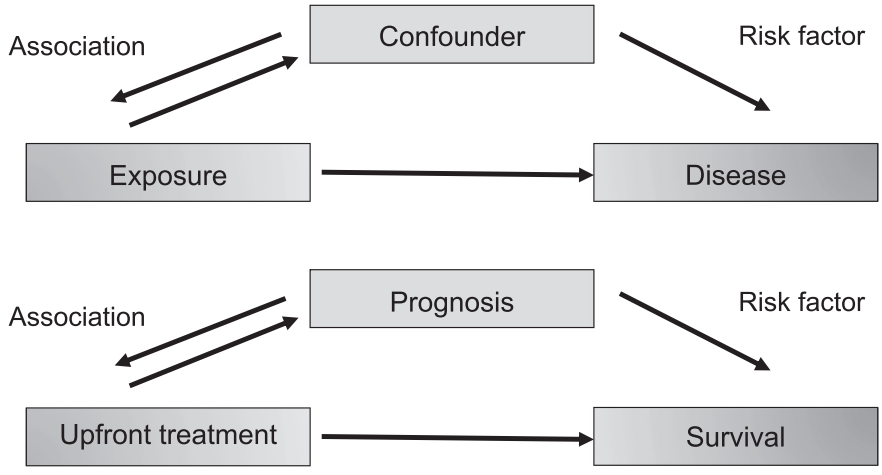

Fig. 1. Demonstrates that upfront treatment is associated with survival. Patient's perceived prognosis at the time of presentation is a confounder as it is associated with both the provider's treatment decision and survival, the outcome of interest.

of confounders, the relationship may not be causal. While the effect of measured confounders on outcomes can be addressed through statistical methods such as multivariable regression analysis and propensity score analysis, these methods cannot adjust for unmeasured confounders.

The problem of unmeasured confounding is common in gynecologic oncology. Important confounders which are unmeasured in many existing data sources include frailty, functional status, and disease burden. Not only are factors which influence treatment selection often absent in existing data sources, measuring these factors is at best challenging, and often infeasible. Provider assessment of patient fitness for a treatment is multifactorial, and rarely based on explicit criteria. Yet such assessment is vital to the selection of approach and extent of surgical resection, as well choice of chemotherapy regimen. Unmeasured confounding may explain discordant results between observational studies and RCTs evaluating the relationship between staging lymphadenectomy and survival in endometrial cancer [5–9] and neoadjuvant chemotherapy for advanced ovarian cancer [10–13].

3. Examples of quasi-experimental methods

3.1. Regression discontinuity design

The regression discontinuity design was introduced in 1960 to study the effect of merit scholarships on college performance [14]. The method has become increasingly common in applied economic research since the late 1990s and has been utilized more recently in health services and epidemiologic research [15–19].

Regression discontinuity designs can estimate causal effects of interventions when the probability of receiving the intervention changes abruptly at a *threshold* value of a randomly distributed continuous variable (called an assignment variable) [17]. When a threshold rule is present, individuals who fall just above and just below the threshold, would be expected to have identical outcomes in the absences of treatment, and are therefore essentially randomized to by the presence of the rule. Regression discontinuity designs are particularly applicable to clinical outcomes research because cut-offs that affect treatment utilization, such as threshold laboratory values or pathologic measurements, are abundant in clinical care.

Thistlethwaite and Campbell's original paper, which investigates the effect of a merit-based college scholarship on the likelihood of obtaining an advanced degree, is a classic example of a regression discontinuity design (Fig. 2). Motivated by limitations of traditional observational studies to estimate unbiased treatment effects, the authors observed that when a merit scholarship is awarded to students who exceed a minimum test score (Fig. 2*A*), student who score just above the cutoff might have scored below by chance alone. Given the probabilistic nature of the precise score, students who score near the threshold are effectively randomized to receiving, or failing to receive, the scholarship by the presence of a threshold. The causal effect of the scholarship on obtaining an advanced degree, can be estimated by calculating the magnitude of the discontinuity in the frequency of advanced degrees

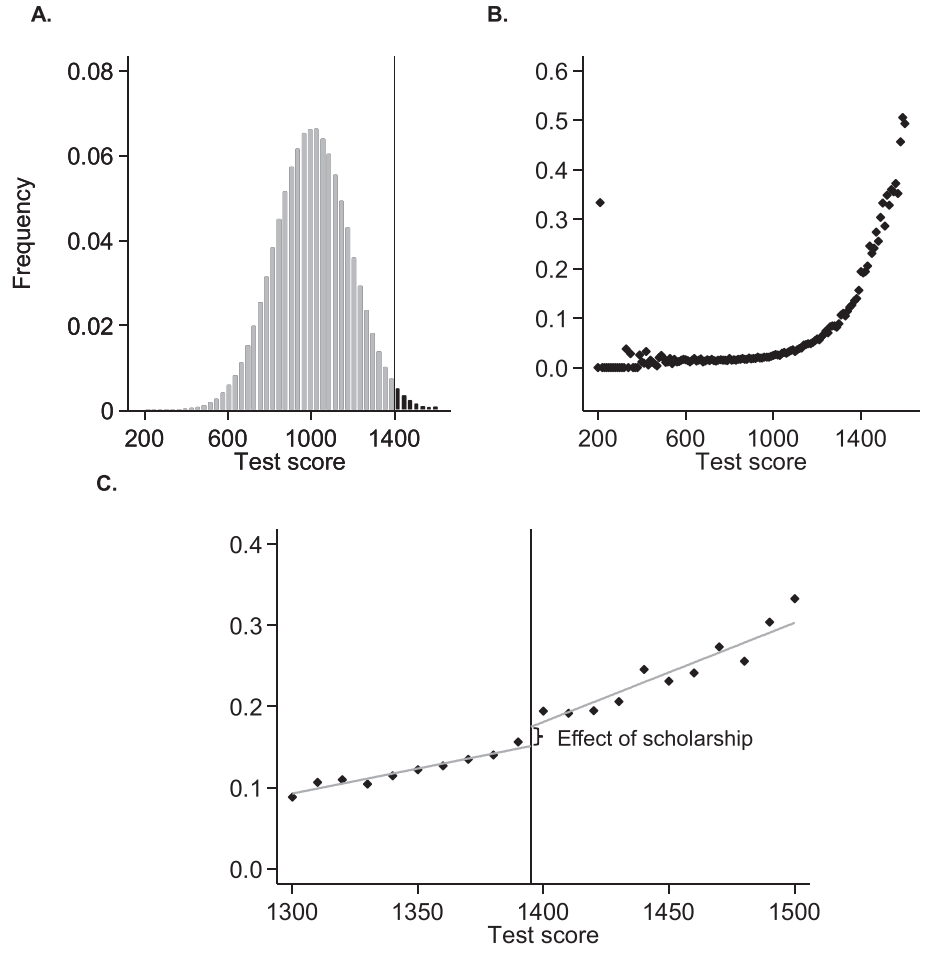

Fig. 2. Graphical representation of a regression discontinuity design to estimate the effect of a merit-based college scholarship (treatment) on obtaining an advanced degree (outcome). Panel A illustrates that a regression discontinuity is possible because receipt of the scholarship is determined by a threshold value of test score (assignment variable). Panel B illustrates a nonlinear relationship between test score and frequency of advanced degree, a subtle discontinuity can be seen at the threshold value of 1400. The magnitude of the discontinuity is estimated in Panel C by fitting a log linear model that estimates the jump associated scoring above the cut-off among students who scored close to the threshold.

between students who score just above and just below the cutoff (Fig. 2B and C).

When a threshold rule does not perfectly predict treatment assignment, but impacts the likelihood of receiving the treatment, as is usually the case in healthcare settings, a fuzzy regression discontinuity design can be used to estimate the impact of the threshold rule on the outcomes. Instrumental variables analysis can be applied to recover the treatment effect among those affected by the rule, known as the complier average causal effect.

Recent studies have used regression discontinuity designs to estimate the effect of antiretroviral therapy on mortality among HIV positive patients [16], evaluate whether a government drug discount program has resulted in intended improvements of care for vulnerable populations in the United States [18], and investigated the impact of adoption of neoadjuvant chemotherapy on ovarian cancer mortality [19]. These diverse studies underscore the utility of regression discontinuity designs to addressing a wide range of clinical and policy questions.

The main advantage of regression discontinuity designs is that in the presence of a threshold, a researcher may draw causal inferences while making relatively modest assumptions [15,20]. As long as a threshold rule is known, the assignment variable is continuous near the threshold, and the potential outcomes are continuous near the threshold, a regression discontinuity can estimate unbiased local treatment effects for subject near the threshold. A major practical limitation to implementing a regression discontinuity design, if that for many interventions of interest there may be no threshold rule, or that an existing rule does not, in practice, impact treatment assignment as expected. Additionally, while regression discontinuities are relatively robust to the presence of unmeasured confounders, a confounder which is discontinuous across the cut-off can result in biased estimates. Finally, strictly speaking, the local effects estimated by regression discontinuity designs may not be generalizable to individuals far from the threshold.

3.2. Difference-in-differences

Difference-in-differences (DiD) is a commonly used statistical technique in observational studies to control for background changes in outcomes that occur with time. The method is commonly used in economics and has become increasingly popular to evaluate outcomes associated with health care policy implementation [21,22]. Evaluating the effect of a particular intervention by comparing outcomes before and after its implementation is only valid when there are no underlying trends in outcomes unrelated to the intervention in question. DiD studies allow for underlying trends in outcomes by adding a control group which does not experience the intervention, but is assumed to experience identical underlying trends [22].

To perform a difference-in-differences analysis the investigator must have access to contemporaneous outcome data for two populations: the treated population, in which an intervention of interest was implemented at a known time, and a comparison population, which is observed during the same period but did not experience the intervention. The difference in outcomes before and after the intervention is estimated in the treated and comparison populations. The effect of the implementation of the intervention can be estimated by comparing the change in outcomes between the treated and comparison groups, the so called difference-in-differences (Fig. 2) [22,23]. If the outcome changed more for the treatment group than the control group, the policy intervention has an impact on the outcome of interest. If there is no relationship between the outcome and the policy implementation, the DiD would equal to 0.

The difference-in-differences design has been used in several studies evaluating the impact of the Affordable Care on insurance status among women with gynecologic cancer. Smith et al. evaluated the effect of the dependent coverage expansion on young women with gynecologic cancer and found that the law increased insurance coverage and early cancer diagnosis [24]. Moss et al. assessed the impact of Medicaid expansion on racial, ethnic and socioeconomic disparities in both expanded and non-expanded states and found that while the proportion of uninsured patients fell across the United States, women living in the highest-poverty areas had the greatest benefit from Medicaid expansion [25]. Difference-in-differences can also be used to evaluate quality improvement initiatives, as it was in a study of a colony-stimulating-factor (CSF) decision support tool, which was found to decrease the use of CSFs with no observed changes in febrile neutropenia among women receiving chemotherapy for breast cancer [26].

There are several limitations of the difference-in-differences method. First, data must exist for the outcome of interest before and after the intervention for both the exposed and unexposed groups. Additionally, data input in cancer-registries and other databases is often delayed making it difficult to study a policy intervention in a timely manner. Second, the exposed and unexposed groups should be similar with the main difference being exposure to the intervention. Therefore, the difference-in-difference methodology relies on the common shock assumption, that any event that occurs during or after the intervention equally affects both groups [22]. Similarly, the parallel trends assumption supposes that in the absence of the intervention, the average outcome of interest would follow parallel paths during the time-period of the study [22,27].

3.3. Interrupted time series

The interrupted time series is a well-known quasi-experimental study design that evaluates the effect of an intervention that has been adopted at a known time [28]. The method relies on the availability of serial outcome measurements from before and after introduction of the intervention in the affected population [28]. Longitudinal outcome data from a period preceding the introduction of the intervention are used to define a pre-existing temporal trend in the outcome, which is extrapolated to predict the expected outcomes in the absence of any intervention (Fig. 3) [29]. The effect of the intervention is estimated by comparing the expected outcomes (called the counterfactual) with observed post-intervention outcomes [28–31].

Interrupted time series is a useful study design for evaluating the effects of policy changes, public health interventions, and rapid changes in treatment paradigms. The design has been used to study of the effects of regionalization of lung cancer surgery on perioperative mortality in Canada [32], of electronic chemotherapy prescribing on medical errors [33], and the adoption of minimally invasive radical hysterectomy on cervical cancer mortality [34].

There are several potential threats to the validity of an interrupted time series. A key assumption is that, in the absence of the intervention, prevailing tends would continue unchanged. This assumption is not valid when outcome trends change in response to factors that coincide with study intervention. Since outcome trends may change over time for reasons unrelated to the intervention, the use of older data to extrapolate expected outcomes may predict an implausible counterfactual [31]. On the other hand, inclusion of too few preintervention data points will decrease the power of the study [35]. Another important consideration is the expected impact of the intervention on the outcome, which should be specified *a priori* and reflected in the impact model chosen for the interrupted time series (Fig. 4) [28,31]. Some interventions are expected to affect outcomes immediately while others may lead to a gradual change in trends; specifying an incorrect impact model can result in false inference. Finally, interrupted time series may be strengthened by including a comparison population that did not experience the intervention under study [36].

3.4. Instrumental variables

Instrumental variable analyses (IVA) have been widely used in economics to estimate causal effect, but is recently becoming more popular in the clinical and biostatistical literature [37]. The method can control

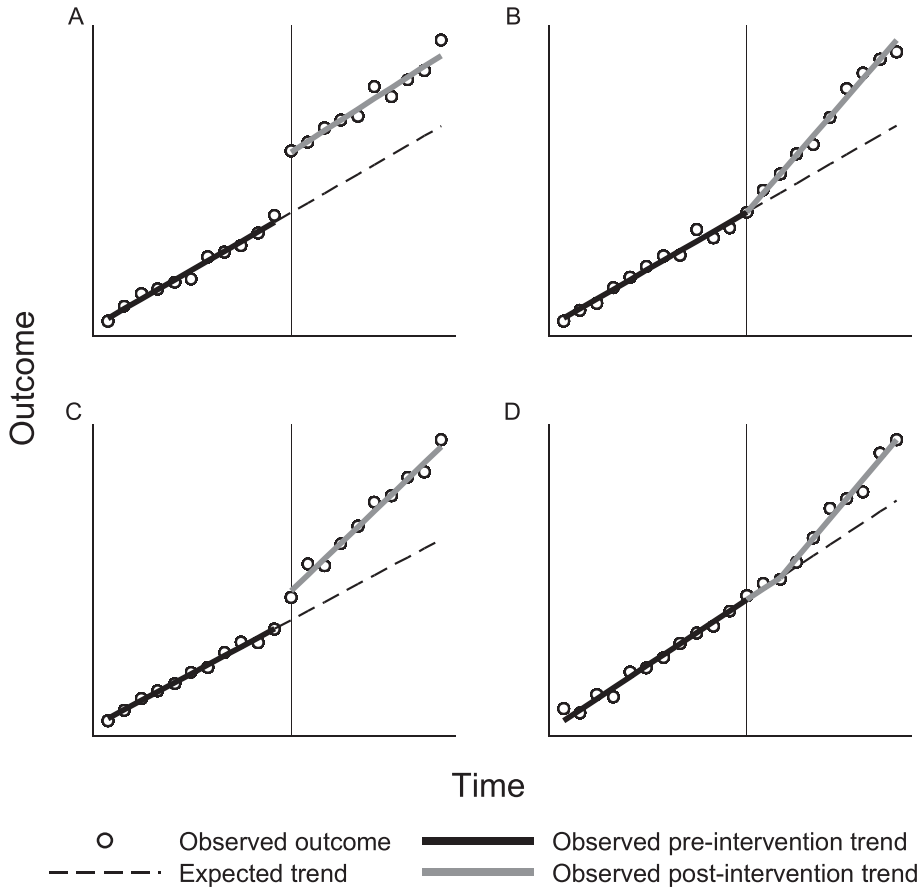

Fig. 3. Difference-in-differences study design. The outcome of interest, plotted on the y-axis, is measured before (pre) and after (post) implantation of an intervention in a treatment group (black circles), and simultaneously in comparison group that is not impacted by the intervention (grey squares). The difference between the change in outcome in the treatment group (*a*) and the change in outcomes in the control group (*b*) is the difference-in-differences, which estimates the causal effect of the intervention on the outcome.

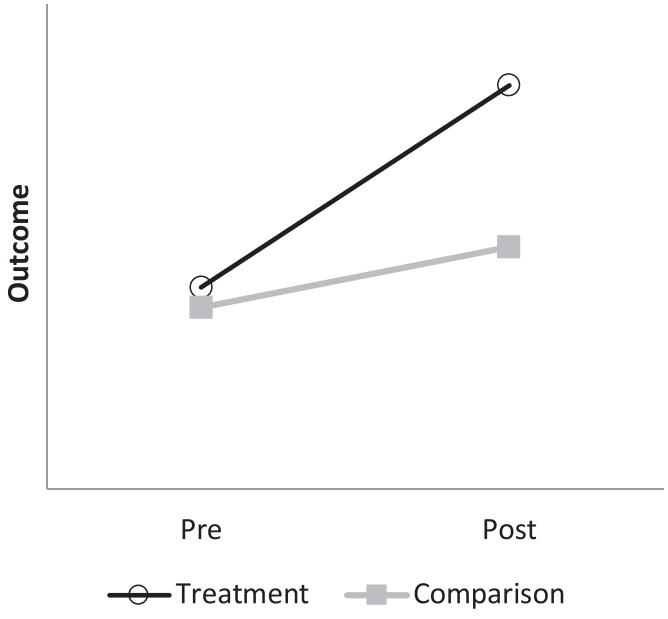

Fig. 4. Impact models for interrupted time series. In all panels the time of the intervention is indicated a vertical line. Pre-intervention outcome trends (solid black lines) are extrapolated to plot the outcome expected in the absence of the intervention (dashed black line). Each panel depicts a different potential response to the intervention such as an immediate rise in the level of the outcome (panel A), a rise in slope of the outcome time trend (panel B), a both a rise in slope and level (panel C), and a delayed rise in slope (panel D).

for both measured and unmeasured confounding, as well as measurement error, in order to estimate the effectiveness of a treatment. The method requires identification of a variable, *the instrument*, that induces substantial variation in the treatment, but has no direct effect on the outcome of interest [37,38]. The causal effect of the treatment on the outcome can be estimated using the variability in the treatment that is due to instrumental variable (Fig. 5) [39].

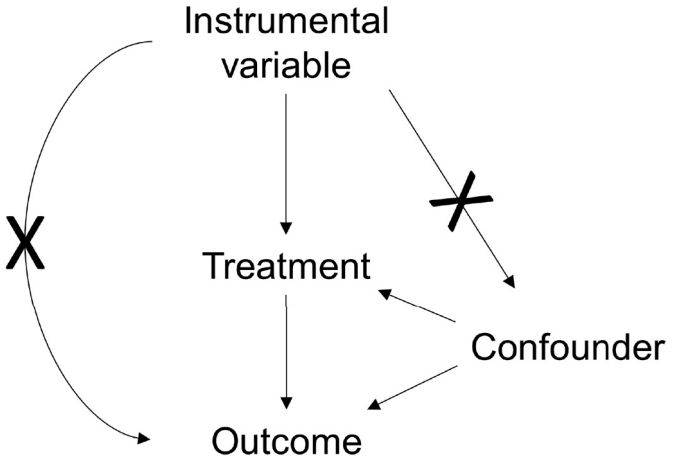

Fig. 5. Diagram of relationship between instrumental variable, treatment, outcome, and confounders. Arrows demonstrate causal relationships. An instrumental variable must affect a subject's likelihood of receiving the treatment but cannot affect the outcome through any other causal pathway.

The instrumental variable must meet the following assumptions: 1) the instrument must have a statistically significant impact on the treatment; 2) it affects the outcome only through its association with the treatment and 3) the relationship between the instrument and the treatment is not confounded by other variables [40,41]. In practice, it is often difficult to identify an instrument, and when a potential instrument is identified, the assumption that the instrument does not affect the outcome, other than through its relationship to the treatment, is difficult to test [39].

A notable example of instrumental variables in clinical outcomes research is a study by McClellan et al. which assess the effect of intensive treatment of acute myocardial infarction on mortality. In this study distance to treatment center is used as the instrument because this variable is unlikely to directly affect acute myocardial infarction related mortality, but does affect the likelihood of catherization [42]. Examples of other instruments used in studies of clinical outcomes include the treatment patterns of a patient's provider, geographic variation in treatment utilization, and economic cost to provider and/or patient [41]. Wright et al. applied IVA methodology using geographic location as an instrument to compare survival between neoadjuvant chemotherapy and primary surgery in a population-based cancer registry. Construction of the instrumental variable demonstrated a strong association between area of residence and primary treatment. The analysis found no difference in survival comparing neoadjuvant chemotherapy to primary surgery for advanced-stage ovarian cancer which was consistent with data from randomized trials [3].

While instrumental variables can be a powerful tool for causal inference in observational data, the method has several limitations. Weak instrumental variables can produce a large standard error which may lead to biased results. Furthermore, the instrumental variable will be imprecise when the sample size is small. In general, power calculations are not straight forward. Likewise, it is advised to only use the instrumental variable method when there is a known variable with a strong correlation to the treatment of interest [39]. The investigator should report the correlation between the instrument and exposure.

4. Selection and implementation of methods

It is important to note that a quasi-experimental design is not available or appropriate for every clinical research question. When unmeasured confounding is not a concern, a quasi-experimental method may have little advantage over a traditional observational study design. However, unmeasured confounding cannot be excluded quantitatively, and both subject matter expertise and empirical evaluation of the data is necessary to assess the potential of unmeasured confounding to bias results [43]. The study designs discussed in this review rely on the presence of real world natural experiments: circumstances in which treatment assignment is, at least in part, dependent on factors beyond patient and provider control. When such a situation occurs, selection of a particular quasi-experiment design depends on the nature of the quasi-experiment, the data available to the investigator, and the plausibility of the designs assumptions for a given research question. For instance, when treatment variation results from threshold rule, a regression discontinuity design is appropriate. Whereas, when treatment variation is the result of a policy change that occurs at a known time (new NCCN guidelines, for example), either an interrupted time series or difference-in-difference might be possible to estimate the policy's effect. The choice between these designs often depends on whether a comparison group or longitudinal outcomes data are available (required for difference-in-differences and interrupted time series respectively). Instrumental variables analysis requires that a plausible instrument can be identified. In general, quasi-experimental methods require larger sample sizes than traditional observational studies; for small samples with rare outcomes, these methods may provide a biased result [44]. An understanding of underlying assumptions, and whether these are concordant with the data being studied is necessary prior to applying a particular quasi-experimental methodology.

5. Conclusion

In the era of big data, well-conducted observational studies are becoming more important than ever. Longitudinal patient databases, cancer registries and data from large electronic health records are providing health service research with new opportunities to perform comparative effectiveness research and examine outcomes in a real-world setting. However, observational studies that fail to consider unmeasured confounders may lead to biased results. Quasi-experimental methods provide investigators the tools to estimate an unbiased treatment effect when confounders cannot be measured. To date, these methods have been underutilized in observational studies in gynecologic oncology. As health information technology spreads and data becomes increasing more available, investigators should consider applying methodologies discussed in this review to strengthen causal inference in observational studies and provide important real-world information which may impact clinical care and health policy.

Author contribution

Haley A. Moss, Alexander Melamed and Jason D. Wright all made substantial contributions to conception and design of the manuscript, participated in drafting the article, revised it critically for intellectual content and gave final approval of the version submitted.

Conflict of interest

The authors declare that there are no conflicts of interest.

CRediT authorship contribution statement

Haley A. Moss: Conceptualization, Project administration, Resources, Writing - original draft, Writing - review & editing. **Alexander Melamed:** Conceptualization, Project administration, Resources, Writing - original draft, Writing - review & editing. **Jason D. Wright:** Conceptualization, Project administration, Resources, Writing - original draft, Writing - review & editing.

References

[1] J.M. Robins, M.A. Hernán, B. Brumback, Marginal structural models and causal inference in epidemiology, Epidemiology 11 (2000) 550–560.
[2] K. Armstrong, Methods in comparative effectiveness research, J. Clin. Oncol. 30 (2012) 4208–4214, https://doi.org/10.1200/JCO.2012.42.2659.
[3] J.D. Wright, C.V. Ananth, J. Tsui, S.A. Glied, W.M. Burke, Y.-S. Lu, A.I. Neugut, T.J. Herzog, D.L. Hershman, Comparative effectiveness of upfront treatment strategies in elderly women with ovarian cancer, Cancer 120 (2014) 1246–1254, https://doi.org/10.1002/cncr.28508.
[4] A.J. Streeter, N.X. Lin, L. Crathorne, M. Haasova, C. Hyde, D. Melzer, W.E. Henley, Adjusting for unmeasured confounding in nonrandomized longitudinal studies: a methodological review, J. Clin. Epidemiol. 87 (2017) 23–34, https://doi.org/10.1016/j.jclinepi.2017.04.022.
[5] L.C. Kilgore, E.E. Partridge, R.D. Alvarez, J.M. Austin, H.M. Shingleton, F. Noojin, W. Conner, Adenocarcinoma of the endometrium: survival comparisons of patients with and without pelvic node sampling, Gynecol. Oncol. 56 (1995) 29–33, https://doi.org/10.1006/gyno.1995.1005.
[6] J.M. Cragun, L.J. Havrilesky, B. Calingaert, I. Synan, A.A. Secord, J.T. Soper, D.L. Clarke-Pearson, A. Berchuck, Retrospective analysis of selective lymphadenectomy in apparent early-stage endometrial cancer, J. Clin. Oncol. 23 (2005) 3668–3675, https://doi.org/10.1200/JCO.2005.04.144.
[7] Y. Todo, H. Kato, M. Kaneuchi, H. Watari, M. Takeda, N. Sakuragi, Survival effect of para-aortic lymphadenectomy in endometrial cancer (SEPAL study): a retrospective cohort analysis, Lancet 375 (2010) 1165–1172, https://doi.org/10.1016/S0140-6736(09)62002-X.
[8] Systematic pelvic lymphadenectomy vs. no lymphadenectomy in early-stage endometrial carcinoma: randomized clinical trial. - PubMed - NCBI (n.d.). https://www.ncbi.nlm.nih.gov/pubmed/19033573 (accessed August 5, 2018).
[9] Efficacy of systematic pelvic lymphadenectomy in endometrial cancer (MRC ASTEC trial): a randomised study. - PubMed - NCBI (n.d.). https://www.ncbi.nlm.nih.gov/pubmed/19070889 (accessed August 5, 2018).

[10] L.A. Meyer, A.M. Cronin, C.C. Sun, K. Bixel, M.A. Bookman, M.C. Cristea, J.J. Griggs, C.F. Levenback, R.A. Burger, G. Mantia-Smaldone, U.A. Matulonis, J.C. Niland, D.M. O'Malley, A.A. Wright, Use and effectiveness of neoadjuvant chemotherapy for treatment of ovarian cancer, J. Clin. Oncol. 34 (2016) 3854–3863, https://doi.org/10.1200/JCO.2016.68.1239.
[11] J.A. Rauh-Hain, A. Melamed, A. Wright, A. Gockley, J.T. Clemmer, J.O. Schorge, M.G. Del Carmen, N.L. Keating, Overall survival following neoadjuvant chemotherapy vs primary cytoreductive surgery in women with epithelial ovarian cancer: analysis of the National Cancer Database, JAMA Oncol. 3 (2017) 76–82, https://doi.org/10.1001/jamaoncol.2016.4411.
[12] I. Vergote, C.G. Tropé, F. Amant, G.B. Kristensen, T. Ehlen, N. Johnson, R.H.M. Verheijen, M.E.L. van der Burg, A.J. Lacave, P.B. Panici, G.G. Kenter, A. Casado, C. Mendiola, C. Coens, L. Verleye, G.C.E. Stuart, S. Pecorelli, N.S. Reed, European organization for research and treatment of cancer-gynaecological cancer group, NCIC Clinical Trials Group, neoadjuvant chemotherapy or primary surgery in stage IIIC or IV ovarian cancer, N. Engl. J. Med. 363 (2010) 943–953, https://doi.org/10.1056/NEJMoa0908806.
[13] S. Kehoe, J. Hook, M. Nankivell, G.C. Jayson, H. Kitchener, T. Lopes, D. Luesley, T. Perren, S. Bannoo, M. Mascarenhas, S. Dobbs, S. Essapen, J. Twigg, J. Herod, G. McCluggage, M. Parmar, A.-M. Swart, Primary chemotherapy versus primary surgery for newly diagnosed advanced ovarian cancer (CHORUS): an open-label, randomised, controlled, non-inferiority trial, Lancet 386 (2015) 249–257, https://doi.org/10.1016/S0140-6736(14)62223-6.
[14] D.L. Thistlethwaite, D.T. Campbell, Regression-discontinuity analysis: an alternative to the ex post facto experiment, J. Educ. Psychol. 51 (1960) 309–317, https://doi.org/10.1037/h0044319.
[15] D.S. Lee, T. Lemieux, Regression discontinuity designs in economics, J. Econ. Lit. 48 (2010) 281–355, https://doi.org/10.1257/jel.48.2.281.
[16] J. Bor, E. Moscoe, P. Mutevedzi, M.-L. Newell, T. Bärnighausen, Regression discontinuity designs in epidemiology: causal inference without randomized trials, Epidemiology 25 (2014) 729–737, https://doi.org/10.1097/EDE.0000000000000138.
[17] A.S. Venkataramani, J. Bor, A.B. Jena, Regression discontinuity designs in healthcare research, BMJ 352 (2016) i1216.
[18] S. Desai, J.M. McWilliams, Consequences of the 340B drug pricing program, N. Engl. J. Med. 378 (2018) 539–548, https://doi.org/10.1056/NEJMsa1706475.
[19] A. Melamed, G. Fink, A.A. Wright, N.L. Keating, A.A. Gockley, M.G. Del Carmen, J.O. Schorge, J.A. Rauh-Hain, Effect of adoption of neoadjuvant chemotherapy for advanced ovarian cancer on all-cause mortality: quasi-experimental study, BMJ 360 (2018) j5463.
[20] E. Moscoe, J. Bor, T. Bärnighausen, Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice, J. Clin. Epidemiol. 68 (2015) 122–133, https://doi.org/10.1016/j.jclinepi.2014.06.021.
[21] A.M. Ryan, J.F. Burgess, J.B. Dimick, Why we should not be indifferent to specification choices for difference-in-differences, Health Serv. Res. 50 (2015) 1211–1235, https://doi.org/10.1111/1475-6773.12270.
[22] J.B. Dimick, A.M. Ryan, Methods for evaluating changes in health care policy: the difference-in-differences approach, JAMA 312 (2014) 2401–2402, https://doi.org/10.1001/jama.2014.16153.
[23] H. Zhou, C. Taber, S. Arcona, Y. Li, Difference-in-differences method in comparative effectiveness research: utility with unbalanced groups, Appl. Health Econ. Health Policy 14 (2016) 419–429, https://doi.org/10.1007/s40258-016-0249-y.
[24] A.J.B. Smith, A.N. Fader, Effects of the affordable care act on young women with gynecologic cancers, Obstet. Gynecol. 131 (2018) 966–976, https://doi.org/10.1097/AOG.0000000000002592.
[25] J.A. Rauh-Hain, W.R. Brewster, K. Behbakht, Society of Gynecologic Oncology 2018 Annual Meeting on Women's Cancer: meeting report, Gynecol. Oncol. 149 (2018) 439–441, https://doi.org/10.1016/j.ygyno.2018.04.569.
[26] A. Agiro, A. Devries, J. Malin, M.J. Fisch, Real-world impact of a decision support tool on colony-stimulating factor use and chemotherapy-induced febrile neutropenia among patients with breast cancer, J. Natl. Compr. Cancer Netw. 16 (2018) 162–169, https://doi.org/10.6004/jnccn.2017.7033.
[27] A. Abadie, Semiparametric difference-in-differences estimators, Rev. Econ. Stud. 72 (2005) 1–19.
[28] J.L. Bernal, S. Cummins, A. Gasparrini, Interrupted time series regression for the evaluation of public health interventions: a tutorial, Int. J. Epidemiol. 46 (2017) 348–355, https://doi.org/10.1093/ije/dyw098.
[29] A.K. Wagner, S.B. Soumerai, F. Zhang, D. Ross-Degnan, Segmented regression analysis of interrupted time series studies in medication use research, J. Clin. Pharm. Ther. 27 (2002) 299–309.
[30] J. Lopez Bernal, S. Soumerai, A. Gasparrini, A methodological framework for model selection in interrupted time series studies, J. Clin. Epidemiol. (2018)https://doi.org/10.1016/j.jclinepi.2018.05.026.
[31] E. Kontopantelis, T. Doran, D.A. Springate, I. Buchan, D. Reeves, Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis, BMJ 350 (2015) h2750, https://doi.org/10.1136/bmj.h2750.
[32] A.M. Bendzsak, N.N. Baxter, G.E. Darling, P.C. Austin, D.R. Urbach, Regionalization and outcomes of lung cancer surgery in Ontario, Canada, J. Clin. Oncol. 35 (2017) 2772–2780, https://doi.org/10.1200/JCO.2016.69.8076.
[33] K. Elsaid, T. Truong, M. Monckeberg, H. McCarthy, J. Butera, C. Collins, Impact of electronic chemotherapy order forms on prescribing errors at an urban medical center: results from an interrupted time-series analysis, Int. J. Qual. Health Care 25 (2013) 656–663, https://doi.org/10.1093/intqhc/mzt067.
[34] L. Chen Melamed, N.L. Keating, M.G. del Carmen, J.D. Wright, J.A. Rauh-Hain, Comparative effectiveness of minimally-invasive staging surgery in women with early-stage cervical cancer, Gynecol. Oncol. 149 (2018) 245–246, https://doi.org/10.1016/j.ygyno.2018.04.553.
[35] F. Zhang, A.K. Wagner, D. Ross-Degnan, Simulation-based power calculation for designing interrupted time series analyses of health policy interventions, J. Clin. Epidemiol. 64 (2011) 1252–1261, https://doi.org/10.1016/j.jclinepi.2011.02.007.
[36] A. Linden, J.L. Adams, Applying a propensity score-based weighting model to interrupted time series data: improving causal inference in programme evaluation, J. Eval. Clin. Pract. 17 (2011) 1231–1238, https://doi.org/10.1111/j.1365-2753.2010.01504.x.
[37] J.P. Newhouse, M. McClellan, Econometrics in outcomes research: the use of instrumental variables, Annu. Rev. Public Health 19 (1998) 17–34.
[38] S. Greenland, An introduction to instrumental variables for epidemiologists, Int. J. Epidemiol. 29 (2000) 722–729, https://doi.org/10.1093/ije/29.4.722.
[39] E.P. Martens, W.R. Pestman, A. de Boer, S.V. Belitser, O.H. Klungel, Instrumental variables: application and limitations, Epidemiology 17 (2006) 260–267, https://doi.org/10.1097/01.ede.0000215160.88317.cb.
[40] M.L. Lousdal, An introduction to instrumental variable assumptions, validation and estimation, Emerg Themes Epidemiol. 15 (2018)https://doi.org/10.1186/s12982-018-0069-7.
[41] J. Hadley, K.R. Yabroff, M.J. Barrett, D.F. Penson, C.S. Saigal, A.L. Potosky, Comparative effectiveness of prostate cancer treatments: evaluating statistical adjustments for confounding in observational data, J. Natl. Cancer Inst. 102 (2010) 1780–1793, https://doi.org/10.1093/jnci/djq393.
[42] M. McClellan, B.J. McNeil, J.P. Newhouse, Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables, JAMA 272 (1994) 859–866.
[43] M.J. Uddin, R.H.H. Groenwold, M.S. Ali, A. de Boer, K.C.B. Roes, M.A.B. Chowdhury, O.H. Klungel, Methods to control for unmeasured confounding in pharmacoepidemiology: an overview, Int. J. Clin. Pharm. 38 (2016) 714–723, https://doi.org/10.1007/s11096-016-0299-0.
[44] M.J. Uddin, R.H.H. Groenwold, A. de Boer, S.V. Belitser, K.C.B. Roes, A.W. Hoes, O.H. Klungel, Performance of instrumental variable methods in cohort and nested case–control studies: a simulation study, Pharmacoepidemiol. Drug Saf. 23 (n.d.) 165–177. doi:https://doi.org/10.1002/pds.3555.