

Toward an Understanding of the Challenges and Opportunities when Studying Emerging Therapies

Christina DeFilippo Mack^{1,2} · Mugdha Gokhale^{2,3}

Published online: 27 October 2016
© Springer International Publishing AG 2016

Abstract

Purpose of Review Emerging therapies and devices may be subject to rapid changes in prescribing or use over short periods, due to recent market launch, approval of a new indication, competitive displacement, or safety warnings. In safety and effectiveness studies of such products, rapid changes in prescribing lead to analytic challenges and opportunities, which may themselves be of interest. This paper discusses how to identify dynamic patterns in dissemination or use, and how to account for and potentially take advantage of these patterns in safety and effectiveness studies.

Recent Findings For every new-to-market drug, the patient population that receives the drug differs from the clinical trial population. Channeling toward or away from certain patient subgroups is expected. In settings where dissemination is dynamic and channeling changes over time are driven by specific patient characteristics, we describe a step-wise approach to understand if there is dynamic dissemination, evaluate how it may impact selection of study cohorts defined by initiation of treatment, and consider different analytic approaches. Calendar time-specific propensity scores (sequential cohort

study) and instrumental variables are highlighted as methods that can take advantage of rapid changes to account for bias and better understand the impacts of treatment dissemination. We illustrate this process through two case studies: (1) a newly indicated therapy for stage 3 colon cancer and (2) a drug recently linked to a safety signal for heart failure.

Summary The authors recommend that researchers systematically evaluate dissemination of products and the impact of policy events on prescribing patterns and choice of alternative treatments. Interim analyses allow for monitoring of patient enrollment, potential changes in channeling over time, and for the purposes of hypothesis generation. Researchers should be cognizant of multiple analytic options available to answer questions of interest.

Keywords Emerging therapies · Drug dissemination · Channeling · Calendar time-specific propensity scores · Instrumental variables · Comparative effectiveness

Introduction

Studies on drugs or devices that experience profound changes in prescribing or use over short period of time warrant additional examination and, potentially, tailored methods. The dynamic nature of these therapies may be driven by new-to-market status or approval for a new indication. Rapid changes in prescribing can also be due to competitive displacement or safety signals that curb or halt use. Analytic methods to account for these patterns in safety and effectiveness studies do not have a “one size fits all” solution; rather, they require individual investigation and contemplation of how best to incorporate information on dissemination and use. We discuss here a recommendation for how to identify dynamic patterns in use or dissemination of a therapy and the methods to

This article is part of the Topical Collection on *Pharmacoepidemiology*

✉ Christina DeFilippo Mack
christina.mack@quintiles.com

Mugdha Gokhale
mugdha.n.gokhale@gsk.com

- ¹ Quintiles, 4820 Emperor Blvd, Durham, NC 27703, USA
- ² Department of Epidemiology, University of North Carolina at Chapel Hill, CB#7435, Chapel Hill, NC 27599, USA
- ³ Epidemiology, Real World Evidence, GlaxoSmithKline, Collegeville, PA 19426, USA

examine, understand, and account for these patterns in safety or effectiveness studies.

Products may experience dynamic dissemination for numerous reasons, some of which are anticipated and others unexpected. Investigation of patterns of use, or retrospective investigation in studies using existing data, is prudent to understand if adjustments in study enrollment or analysis are necessary. Acknowledging that in a sense, all drugs arriving to market are subject to this and at minimum, a light investigation is routinely done, specific settings drive particular attention to this:

- A new drug targeted toward a disease or indication with no treatment, for example Viagra's market launch in 1998 [1]
- Disease areas such as oncology and emerging/infectious disease where treatment markets are more commonly innovative and receptive to new interventions [2]
- Settings where the current direct competitor has safety questions or new randomized trial evidence emerges that indicates lack of effectiveness, and a shift between treatment alternatives for the same indication is expected
- Release of safety data prompting speculation or issuance of a black box warning
- Studies mandated as a condition of approval such as post-approval safety studies, particularly when these studies are comparative and differences between the populations are expected

For every new-to-market drug, the precise characteristics of the patient population that will receive the drug are unknown; we only know that this population is going to be different than patients in the clinical trials. Channeling of the therapy based on underlying patient characteristics, either toward or away from certain patients, is expected. The answers to what if anything is differentially dictating patient receipt and how this changes throughout the course of the drug's lifetime needs to be explored, identified, and accounted for in analytics to achieve robust results. In addition, such trends offer analytic opportunities that should be explored. Describing these trends is a critical part of understanding a new drug, with more emphasis required when dynamic dissemination or changes in channeling are identified.

In settings where dissemination is dynamic and channeling changes over time is driven by specific patient characteristics, calendar time-specific propensity scores (sequential cohort study) and instrumental variables are tools that can be used together or separately [3–6]. Propensity scores estimate the probability that a patient receives one treatment over another, as a function of patient characteristics at the time of the treatment decision [7, 8]. The predictive values assigned within the propensity score provides insight into the probability of treatment receipt based on each patient characteristic; in settings of

rapid dissemination, estimating the propensity score in calendar blocks gives transparency into whether this changes over the study timeframe. Instrumental variables, in pharmacoepidemiology, are variables that directly affect treatment receipt but are not related to the study outcome, either directly or indirectly [9]. This effectively pseudo-randomizes the population, allowing for a natural experiment that generates variation in the exposure as in randomized assignment. If a change in prescribing or use occurs rapidly over time across the patient population, time may act as an instrument for treatment receipt.

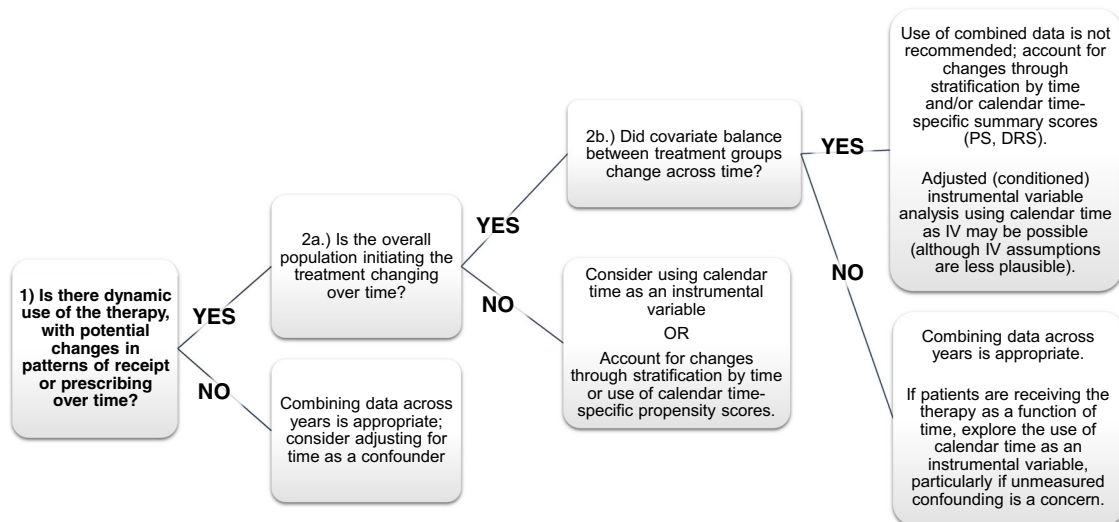
Here, we discuss some of the questions to consider and recommendations (non-exhaustive) in designing and analyzing studies with dynamic treatment settings (Fig. 1). We will use two case studies to illustrate these recommendations:

Case Study 1: Oxaliplatin vs. 5-Fluorouracil for Stage 3 Colon Cancer

Oncologic drugs are often disruptive due to the prognosis of patients and the continuing search for better treatments. Colon cancer is the second leading cause of cancer mortality in the USA; stage III colon cancer has a 3-year survival rate of 36 % [10]. Oxaliplatin, a drug approved for stage IV colon cancer, experienced rapid uptake among stage III patients, displacing 5-fluorouracil (5-FU) monotherapy as the standard of care over a very short time period. From 2003 to 2007, oxaliplatin, in a combination regimen with 5-fluorouracil (5-FU) and folinic acid, rapidly disseminated among stage III colon cancer patients to replace 5-FU monotherapy as the standard of care based on 2003 efficacy results from the MOSAIC trial [2, 11] and subsequent US Food and Drug Administration (FDA) approval in November 2004. This rapid uptake among patients, many of whom were elderly and underrepresented in clinical trials, contributed to a setting where a patient's likelihood of receiving oxaliplatin vs. 5-FU was determined predominately by the time in which their treatment decisions were being made (before or after FDA approval of oxaliplatin for stage III), rather than based on patient characteristics. At the same time, as the drug remained on the market, prescribing patterns based on patient characteristics were modified. Taken together, these prescribing paradigms between 2003 and 2005 render oxaliplatin an interesting example of drug dissemination and the need for careful consideration toward analytic methods as a result.

Case Study 2: Dipeptidyl Peptidase 4 Inhibitors (DPP-4i) vs. Thiazolidinediones (TZD) and Heart Failure

The antihyperglycemic drug market has seen introduction of several new drugs in the last decade along with emerging safety



*Subject to assumptions, including no unmeasured confounding for adjusted analyses.

Fig. 1 Questions and analytic recommendations* when dynamic patterns of use are expected

concerns about thiazolidinediones (TZD). In 2010, the FDA severely restricted the use of rosiglitazone due to concerns about increased risk of myocardial infarction and issued a drug safety communication about the increased risk of bladder cancer with pioglitazone. According to the recent market data, the newly approved drugs, particularly the dipeptidyl peptidase 4 inhibitors (DPP-4i) class, quickly gained significant market share and was the most commonly prescribed new drug class by 2012 [12]. On the other hand, the use of TZD decreased by 2012; in particular, pioglitazone use reached only half its peak use in 2008 and rosiglitazone decreased to negligible use. These dynamic patterns created a unique setting where DPP-4i and TZD can be compared to each other in a conventional new user active comparator cohort study (hereafter referred to as active comparator approach); but the crossing of drug initiation curves also provided an opportunity to use calendar time as an instrumental variable to compare the risk of an outcome. The key point in this study was that heart failure was used as a positive control outcome and the researchers could compare the results from the two analytic methods knowing that DPP-4i would have a lesser risk of heart failure relative to TZDs.

1. Investigate Dissemination of Drug Across Target or Study Population

We know that certain patients will be more likely to receive the drug than others. Unknowns include the following:

- Which patients experience channeling?
- Which characteristics cause them to receive or not receive the drug?
- Will this change over time?

Time is the first critical variable of focus here, both around therapy milestones (e.g., FDA approval, release of safety warnings or post-approval study results, broad reimbursement decisions) as well as regular points in calendar time. The goal is to see if utilization has abruptly changed. In retrospective studies, this can be done in early stages of analysis to see if patterns changed. In prospective studies, examination of use and prescribing patterns should be planned for and done annually, biannually or more frequently based on events in the drug's lifecycle.

Case Study 1 In a retrospective analysis of use among stage 3 colon cancer, restricting only to patients who received either oxaliplatin or 5-FU, the standard of care, 100 % of study patients received 5-FU from March 2003–June 2003. As results from the MOSAIC trial became known [11, 13], use of oxaliplatin increased, and by February 2005, 3 months after FDA approval, oxaliplatin use was (with one exception) consistently above 50 % (Fig. 2). By 2.5 years after FDA approval (post-May 2007), over 60 % of patients in the study population had received oxaliplatin each month for a full year, indicating that it had become the standard of care [14]. Thus, it was clear that dissemination was occurring rapidly.

Case Study 2 In a retrospective analysis of Medicare enrollees initiating oral second-line antidiabetic drugs between 2008 and 2013, the proportion of sulfonylurea initiators was somewhat constant at ~70 % and the initiation of second-line diabetes therapy with DPP-4i increased while TZD decreased. The two drug initiation curves crossed in 2010 around the time of FDA warnings about the safety of TZD (Fig. 3). Due to such dynamic patterns of initiation over a relatively short period of time, the authors investigated whether the implementation of a

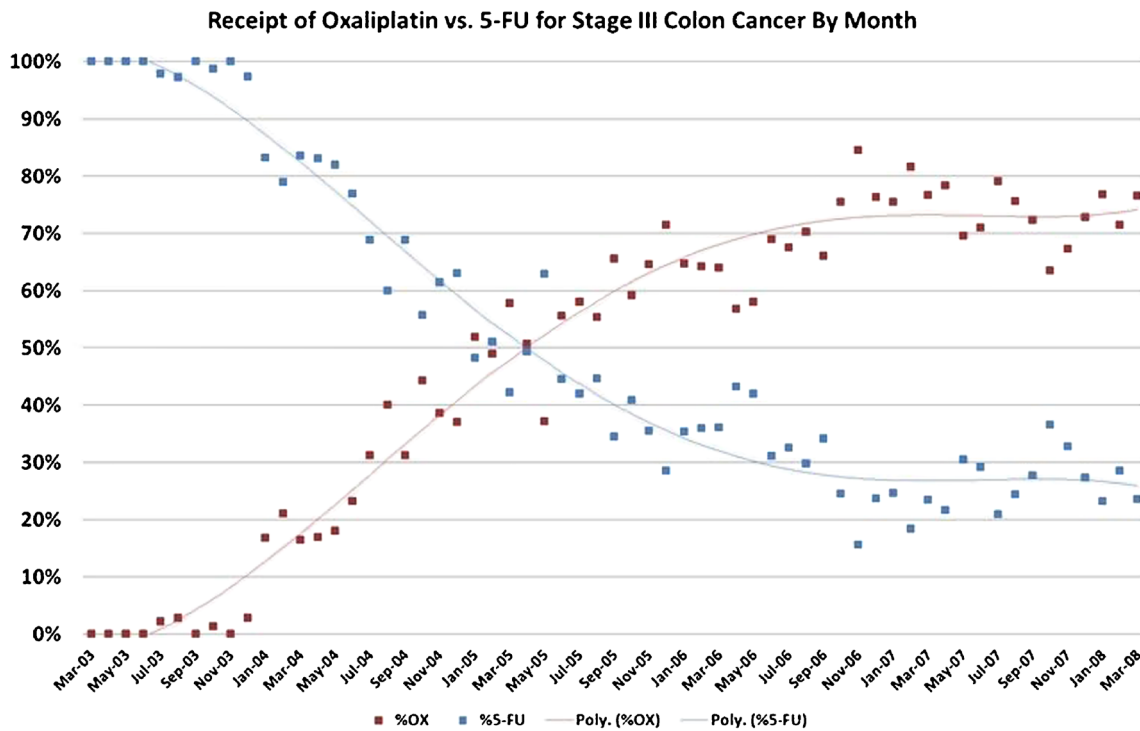


Fig. 2 Receipt of oxaliplatin vs. 5-FU for stage III colon cancer by month (Adapted from: Copyright © 2015 Wolters Kluwer Health, Inc.; Mack et al. [14])

new user cohort study design using combined data from all years would be affected. Specifically, they examined whether the overall study population selected in the earlier years was different than the population initiating the same drug class in later years (2a). Next they looked at whether the covariate balance between DPP-4i versus TZD changed (2b).

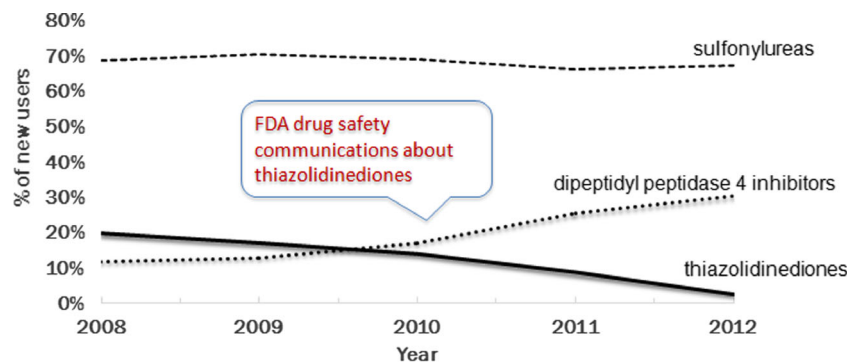
2a. Is the Overall Population Initiating the Treatment Changing over Time, Therefore Changing the Population Selected for the Study?

Changes in treatment initiation across the target population may be due to changes in drug indication, insurance coverage, or other events during the course of a study. This can be investigated by looking at the absolute change in the prevalence of

covariates over time for the study population and should also be considered qualitatively using expert knowledge of the therapy lifecycle on a global level.

Case Study 1 The expectation among this group of patients is that stage III colon cancer patients who were eligible for chemotherapy would have either received 5-FU or oxaliplatin across the course of the study timeframe. Sensitivity analyses of excluded patients suggested that the overall population selected for the study did not change; the same patients appeared to be receiving different drugs across years. Therefore, the study population overall was understood to be consistent, although the treatment received within these patients could possible change over time. The next step was to investigate whether the population receiving the drug changed over time based on patient characteristics (2b).

Fig. 3 Receipt of second-line diabetes drugs sulfonylureas, dipeptidyl peptidase 4 inhibitors and thiazolidinediones across years (Gokhale et al. [15])



Case Study 2 In this case, TZD are contraindicated in patients with HF. Therefore, a study examining cardiovascular outcomes comparing DPP vs TZD in a population including patients with existing HF would be intractably confounded by contraindication. To avoid that, patients with HF or related conditions (for example, use of loop diuretics) at baseline were excluded. In the restricted population, DPP and TZD are expected to be in equipoise supporting the use of an active comparator cohort study. Examination of covariates in the overall population across time revealed that in spite of the dynamic treatment patterns, no difference was found in the characteristics of patients selected in early versus later years of this restricted population.

2b. Did Covariate Balance Between Treatment Groups Change Across Time?

Examination of the absolute change of covariate prevalence or average standardized absolute mean difference (ASAMD [16]) between treatment cohorts informs the best way to proceed at this step; as a general rule, less than 10 % difference indicates “balance” between the groups. Alternatively or in addition to investigating each variable independently, logistic propensity score models can be fitted within specific time points, and the coefficients within these models can be used to investigate the conditional relationship that each variable has with treatment to determine whether the odds for treatment receipt for a given covariate changes over time.

Case Study 1 To understand this, the authors investigated variables in regular half and full year time points, anchored at the presentation of MOSAIC study results in May 2003. They found that patients with diabetes became less likely to receive oxaliplatin after the first year of the study, when diabetes did not impact treatment receipt. A potential explanation for this is that as physicians gained experience with the drug, they observed that the known neurotoxic side effect of oxaliplatin disproportionately affects patients with diabetes, thus discouraging them from prescribing oxaliplatin to diabetic patients [6]. Patients that lived in higher income areas appeared to have increased access to oxaliplatin in its first year on the market, despite the fact that all patients in the study were covered exclusively by Medicare; this difference lessened after FDA approval but did not disappear. Consistent channeling of oxaliplatin away from the oldest patients in the study suggested that unmeasured variables such as frailty or age discrimination may have become increasingly relevant factors in provider and/or patient decision-making (Fig. 4). These findings indicate that the effects of covariates on treatment vary over time. This can be accounted for by constructing propensity scores in separate time blocks or by including time interactions within an overall propensity score.

While the time-stratified propensity score approach deals with the imbalances in measured factors over time, it does not address balances in unmeasured confounding such as unmeasured frailty, which is a particular concern in this population of elderly patients. For this reason, the authors also performed an instrumental variable analysis (discussed in detail below).

Case Study 2 Closer examination of the covariates across time revealed that the covariate balance between DPP-4i and TZD did not meaningfully change throughout the study period. For all covariates, the odds ratios indicating channeling for DPP-4i versus TZD were very similar across time changing by <5 %. Therefore, the setting did not warrant the use of calendar time-specific propensity scores and the researchers used combined data with overall propensity scores to control for confounding. The authors also explored the use of calendar time as an instrumental variable and found that both analytic approaches performed equally well in this setting.

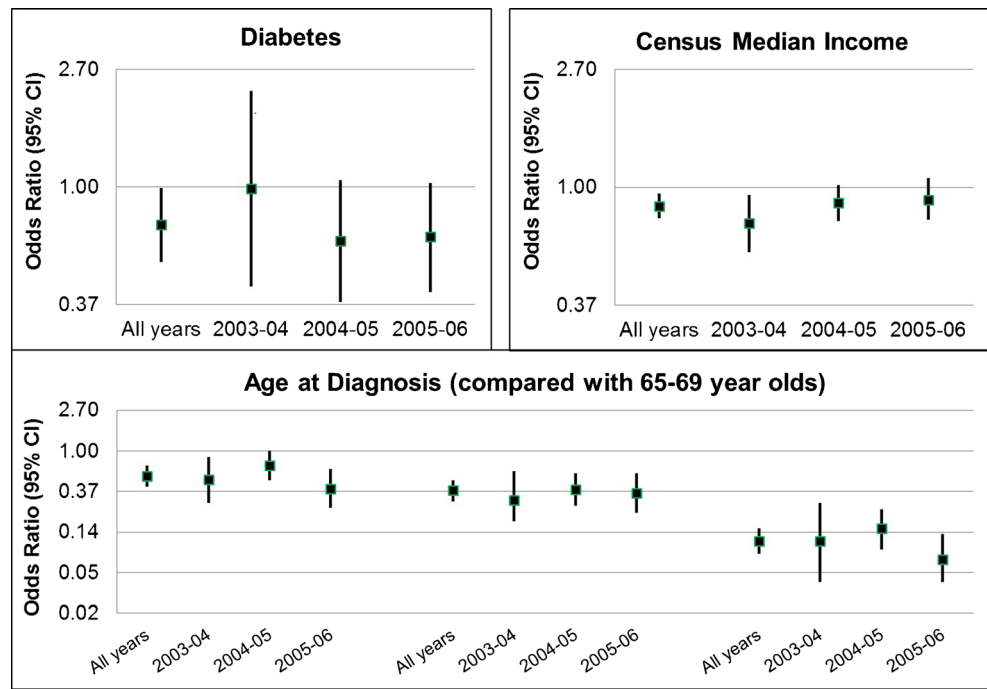
Analytic Recommendations

If patient receipt of treatment differs by year, calendar time-specific propensity scores and instrumental variables are among methods to consider as primary analyses, after careful assessment of relevant assumptions (for a full discussion of IV assumptions, see Swanson et al. [17]). These methods are also useful as secondary analytic approaches in settings where combined data can be used, as unmeasured confounding is often present in some degree and both methods strengthen the evidence of the full study by adding another data point to the estimates. Our case examples illustrate two ways to approach this, as they present considerations within different settings (new indication vs. safety signal). The examples also differ with respect to the potential for unmeasured confounding, as for case study 1 this was of high concern, whereas unmeasured confounding was less of a risk in the case study 2 comparison between DPP-4i and TZD.

Calendar Time-Specific Propensity Scores

If channeling toward or away from specific patients changes over time due to dissemination patterns, a calendar time-specific propensity score is appropriate. Propensity score methods assume that all confounders are accurately measured and the specified model of treatment receipt is correct. Often, propensity scores are estimated across a full study time period with calendar time in the model, which assumes a constant effect of patient characteristics on treatment choice over multiple study years. In a dynamic setting, this will not produce an accurate model, as a patient characteristic that was once

Fig. 4 Changes in channeling by covariate over study time periods (adjusted OR, 95 % CI) comparing receipt of oxaliplatin with 5-FU (Republished with permission of John Wiley and Sons, Pharmacoepidemiol Drug Safety, Mack CD, Glynn RJ, Brookhart MA, Carpenter WR, Meyer AM, Sandler RS, Stürmer T, 2013;22(8):10.1002/pds.3386; permission conveyed through Copyright Clearance Center, Inc.)



associated with treatment selection may become less relevant over time, or vice versa.

Case Study 1 The authors constructed a calendar time-specific propensity score to examine channeling within each time period and to reduce bias by accounting for changes in predictors of treatment receipt over the study time period. This method allowed the influence of each covariate on the propensity for treatment receipt to be non-uniform across the study and thus take into account changes in channeling by specific patient characteristics as oxaliplatin rapidly disseminated, replacing 5-FU as the standard of care [6]. This improved covariate balance within calendar time strata and likely resulted in enhanced confounding control.

Case Study 2 Because the channeling between DPP-4i versus TZD did not change across time for any of the covariates, use of overall propensity scores was deemed appropriate. Because calendar time strongly affected treatment initiation in this setting, the researchers did not include time in the PS estimation models as adjusting for an instrument within a propensity score can bias estimates and negatively impact precision [18].

Instrumental Variables

If the reason patients are getting the new drug is the calendar time that they were diagnosed, separate from other factors, time may act as an instrument for treatment receipt. This assumes that time, as an instrument, (1) strongly affects or is associated with the exposure, (2) is related to the outcome

only through its association with the exposure, and (3) is unrelated to patient risk factors for the outcome. While an instrumental variable analysis depends on these often difficult-to-meet assumptions [19], it does not rely on the assumption of unmeasured confounding, which provides benefits in a number of situations where difficult-to-quantify variables such as frailty or patient decision-making are strong confounders.

To investigate whether calendar time is strongly tied to treatment receipt, examine the by month or by quarter dissemination between the key therapies, with particular attention to product milestones such as regulatory approval, publication of safety issues, or reimbursement changes (Figs. 2 and 3). Crossovers in usage patterns suggest that time may directly impact treatment receipt (IV assumption 1), at which time, the other assumptions can be considered.

Case Study 1 Calendar time strongly impacted treatment receipt, as oxaliplatin treatment rates were 11 % in the early arms of the instrument (January 2003–September 2004) vs. 65 % in the late arm (March 2005–May 2007). This yielded an instrument strength, also called the compliance percentage or scaling factor [17], of 54 %, which is considered to be a very strong instrument [14]. Given that the IV well met assumption 1, the authors examined and tested the other assumptions to the extent possible using an examination of patient characteristics across the instrument and treatment, expert knowledge around potential changes that might impact the outcome of mortality across time, and falsification tests [17, 20] and decided that an instrumental variable analysis was possible and would provide benefits in not relying on the assumption of unmeasured confounding.

Instrumental variable results were consistent with propensity score-adjusted results in the conclusion that oxaliplatin was protective over 5-FU within a broad, older population. However, the magnitude of the instrumental variable risk difference was slightly greater than that observed through propensity score methods. Also notable here is that instrumental variable estimates apply to a different population than the target population in a propensity score analysis [14, 17].

Case Study 2 As demonstrated by the crossing of the drug initiation curves (Fig. 3), calendar time strongly influenced initiation of DPP-4i versus TZD. This created a “natural experiment” suitable for IV methods because the allocation of the exposure was similar to that of a randomized experiment. The DPP-4i treatment rates in the “post” (October 2010–December 2013) and “pre” (January 2008–May 2010) periods were 77 % and 37 %, respectively, leading to a compliance percentage of 40 %, indicating that calendar time is again a very strong instrument. After evaluating the other assumptions as described above, the use of IV methods was possible. Both the IV and active comparator propensity score approaches indicated lesser risk of HF hospitalizations with DPP-4i vs. TZD initiators.

Discussion

We provide here a guideline for studying dynamic therapies within the context of two specific examples. This is intended to highlight issues around rapid treatment changes; it is not intended as an exhaustive review of this complex and broad topic.

The cases described use both a propensity score and instrumental variable approach, which produced a range of estimates based on different assumptions. While the true magnitude of the estimates of effectiveness cannot be confirmed due to the inability to empirically test assumptions, the different methods provide a useful range of values and, in their similarity to overall conclusions of effectiveness, lend credibility to the study conclusions.

In both case studies 1 and 2, the instrumental variable estimate was larger than the propensity score estimate. This may be due to the scaling of the IV, as scaling may magnify biases (if any) in the unscaled estimates [19]. A recent paper by Jackson and colleagues suggests that scaling can have implications in bias due to imbalance of covariates across the levels of the IV [21]. Specifically, it can be misleading to directly compare covariate balance across the levels of the “treatment received” and the unscaled covariate balance across the levels of the IV; even though commonly recommended, the later underestimates the imbalance. The authors suggest examining the scaled covariate balance, i.e., the prevalence difference across IV levels multiplied by the scaling factor, instead.

[22] Both the above points are more of a concern with weak instruments (less compliance) because the scaling factor becomes increasingly large as the instrument gets weak. Given this and other limitations [19, 22], there are still merits to instrumental variable analyses in settings similar to case studies 1 and 2 that result in very strong IVs.

If time is deemed to be an instrument, it is important that it is not included as an adjustment variable in the propensity score; this will decrease precision [17] and increase bias in the presence of unmeasured confounding [23–26]. That said, Myers et al. argue that it might be prudent to control for variables, such as calendar time in this instance, unless they undoubtedly fulfill all the IV conditions; this is often hard to demonstrate empirically. Such variables may be related to changes in coding or incidence of the outcome rather than serving as an instrument of treatment [18]. In cases where both the CTS PS and IV method is used, calendar time can be used to stratify but not be included as an adjustment variable in the PS model.

Sensitivity analyses around time points used within the calendar time-specific propensity score or for the instrumental variable may be warranted. For the IV analyses in particular, it is valuable to investigate different cutpoints, considering the IV assumptions and calculating the compliance percentage around each option to optimize the balance between the strength and validity of the instrument. This was done in both of the case studies above.

Disease risk scores are an alternative method that can also be useful in dynamic settings, particularly when prescriber preferences are continually evolving [27]. Unlike the propensity score which summarizes covariate associations with treatment assignment, the disease risk score summarizes covariate associations with potential outcomes [28]. Because summary measures of disease risk reflect what may be a biologic relationship to the outcome, they are more likely to be stable over time compared to covariate relationships with treatment assignment. Disease risk scores may hold particular advantages in the early monitoring of newly introduced therapies where treatment patterns are rapidly changing and exposure is often infrequent [27, 29]. The benefits and challenges of applying risk scores when evaluating evolving drug therapies are discussed in more detail elsewhere in this issue [30].

We recommend that drug researchers evaluate dissemination of dynamic products with frequent interim analyses which allow for monitoring of patient enrollment, potential channeling and changes in channeling, and for the purposes of hypothesis generation. When the propensity for treatment receipt may change for one or more patient characteristic, it is important to examine whether channeling is occurring in real time. This agile approach to analytics, allowing for responsiveness to observed patterns or events, is warranted in these settings [31]. One method to do this is to repeatedly generate a time-specific propensity score and compare the propensity for

treatment receipt across time periods by covariate. This information can inform decisions about study design going forward. If receipt changes drastically such that a specific subgroup becomes the main recipient of the new treatment, the propensity score model would give insight into the individual contribution of covariates to the prediction of treatment changes over time [6]. Automated options for generating propensity scores may be faulty in these settings, as they may not be enough to detect and account for dissemination patterns. Rather, therapies that are rapidly disseminating require a manual look and thoughtful consideration of patterns seen in data, as well as knowledge of clinical use and patient population.

Conclusions

It is important to examine the impact of time throughout studies of all treatments, and particularly those that are candidates for dynamic dissemination. This has become an even more relevant issue with the increase in post-approval effectiveness studies and strong practices for safety monitoring of existing treatments. Comparative studies in particular may be more impacted by time trends, since clinical equipoise between treatments may change if there is a safety signal (illustrated in case study 2, above); the availability of a second option then allows rapid changes in use. There are a number of methods available, both for understanding, accounting for and taking advantage of these dissemination patterns in study design and analysis. While excellent literature exists individually on different methodologic options, in this paper, we recommend diligence to this process as well as provide a tactical explanation of approach.

Compliance with Ethical Standards

Conflict of Interest Mugdha Gokhale is a full time employee of GlaxoSmithKline. Christina Mack is a full time employee of Quintiles, a clinical research organization which provides services to pharmaceutical and medical device companies.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

1. Padma-Nathan H. Sildenafil citrate (Viagra) treatment for erectile dysfunction: an updated profile of response and effectiveness. *Int J Impot Res.* 2006;18(5):423–31.
2. Buzdar A, Macahilig C. How rapidly do oncologists respond to clinical trial data? *Oncologist.* 2005;10(1):15–21.
3. Seeger JD, Walker AM, Williams PL, Saperia GM, Sacks FM. A propensity score-matched cohort study of the effect of statins, mainly fluvastatin, on the occurrence of acute myocardial infarction. *Am J Cardiol.* 2003;92:1447–51.
4. Seeger JD, Williams PL, Walker AM. An application of propensity score matching using claims data. *Pharmacoepidemiol Drug Saf.* 2005;14:465–76.
5. Schneeweiss S, Gagne JJ, Glynn RJ, Ruhl M, Rassen JA. Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. *Clin Pharmacol Ther.* 2011;90(6):777–90 Epub 2011 Nov 2.
6. Mack CD, Glynn RJ, Brookhart MA, Carpenter WR, Meyer AM, Sandler RS, Stürmer T. Calendar time-specific propensity scores and comparative effectiveness research for stage III colon cancer chemotherapy. *Pharmacoepidemiol Drug Saf.* 2013;22(8). doi:10.1002/pds.3386.
7. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika.* 1983;70:41–55.
8. Stürmer T, Schneeweiss S, Brookhart MA, Rothman KJ, Avorn J, Glynn RJ. Analytic strategies to adjust confounding using exposure propensity scores and disease risk scores: nonsteroidal antiinflammatory drugs and short-term mortality in the elderly. *Am J Epidemiol.* 2005;161(9):891–8.
9. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiol Drug Saf.* 2010;19(6):537–54 Review.
10. Altekruse SF, Kosary CL, Krapcho M, et al., editors. SEER cancer statistics review, 1975–2007. Bethesda, MD: National Cancer Institute ; 2010. Available at: http://seer.cancer.gov/csr/1975_2007/
11. André T, Boni C, Mounedji-Boudiaf L, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N Engl J Med.* 2004;350:2343–51.
12. Hampp C, Borders-Hemphill V, Moeny DG, Wysowski DK. Use of antidiabetic drugs in the US, 2003. *Diabetes Care.* 2014;37(5):1367–74.
13. National Cancer Institute. Highlights from ASCO 2003. May 31, 2003. Available at: <http://www.cancer.gov/asco2003/highlights>
14. Mack CD, Brookhart MA, Glynn RJ, Meyer AM, Carpenter WR, Sandler RS, Stürmer T. Comparative effectiveness of oxaliplatin vs. 5-fluorouracil in older adults: an instrumental variable analysis. *Epidemiology.* 2015;26(5):690–9.
15. Gokhale, M. N., Buse, J. B., Marquis, M. A., & Sturmer, T. (2015, September). Patterns of Antihyperglycemic Drug Use and Effect on Implementation of a New-User Active Comparator Study. In *Pharmacoepidemiology and drug safety* (Vol. 24, pp. 8-8). 111 River St, Hoboken 07030-5774, NJ USA: WILEY-BLACKWELL.
16. Ellis AR, Dusetzina SB, Hansen RA, Gaynes BN, Farley JF, Stürmer T. Confounding control in a nonexperimental study of STAR*D data: logistic regression balanced covariates better than boosted CART. *Ann Epidemiol.* 2013;23(4):204–9.
17. Swanson SA, Hernán MA. Commentary: how to report instrumental variable analyses (suggestions welcome). *Epidemiology.* 2013;24(3):370–4.
18. Myers JA, Rassen JA, Gagne JJ, Huybrechts KF, Schneeweiss S, Rothman KJ, et al. Effects of adjusting for instrumental variables on bias and precision of effect estimates. *Am J Epidemiol.* 2011;174(11):1213–22.
19. Hernan MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology.* 2006;17(4):360–72.
20. Glymour MM, Tchetgen EJ, Robins JM. Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. *Am J Epidemiol.* 2012;175(4):332–9.

21. Jackson JW, Swanson SA. Toward a clearer portrayal of confounding bias in instrumental variable applications. *Epidemiology* (Cambridge, Mass). 2015;26(4):498.
22. Swanson SA, Hernán MA. Think globally, act globally: an epidemiologist's perspective on instrumental variable estimation. *Stat Sci*. 2014;29(3):371–4. doi:10.1214/14-STS491.
23. Patrick AR, Schneeweiss S, Brookhart MA, Glynn RJ, Rothman KJ, Avorn J, et al. The implications of propensity score variable selection strategies in pharmacoepidemiology: an empirical illustration. *Pharmacoepidemiol Drug Saf*. 2011;20(6):551–9.
24. D'Agostino Jr RB, D'Agostino Sr RB. Estimating treatment effects using observational data. *JAMA*. 2007;297(3):314–6.
25. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. *Am J Epidemiol*. 2006;163:1149–56.
26. Pearl J. Invited commentary: understanding bias amplification. *Am J Epidemiol*. 2011;174:1223–7.
27. Glynn RJ, Gagne JJ, Schneeweiss S. Role of disease risk scores in comparative effectiveness research with emerging therapies. *Pharmacoepidemiol Drug Saf*. 2012;21(Suppl 2):138–47.
28. Hansen BB. The prognostic analogue of the propensity score. *Biometrika*. 2008;95(2):481–8.
29. Wyss R, Ellis AR, Brookhart MA, et al. Matching on the disease risk score in comparative effectiveness research of new treatments. *Pharmacoepidemiol Drug Saf*. 2015;24(9):951–61.
30. Wyss R, Glynn RJ, Gagne JJ. A review of disease risk scores and their application in pharmacoepidemiology. *Current Epidemiology Reports*. 2016. In Press.
31. Mack CD, Parmenter L, Velentgas P, Franke K, Jablonski R, Dreyer N. Agile analytics—delivering actionable results for observational studies through rapid evaluation processes. Washington DC: ISPOR 21st Annual International Meeting; 2016. p. 21–25.