Drug and Alcohol Dependence 169 (2016) 180–189

Contents lists available at ScienceDirect

Drug and Alcohol Dependence

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

Review

Selection criteria limit generalizability of smoking pharmacotherapy studies differentially across clinical trials and laboratory studies: A systematic review on varenicline

Courtney A. Motschman[a], Julie C. Gass[a], Jennifer M. Wray[a,b], Lisa J. Germeroth[a], Nicolas J. Schlienz[a,c], Diana A. Munoz[a], Faith E. Moore[a,d], Jessica D. Rhodes[a,e], Larry W. Hawk[a], Stephen T. Tiffany[a,*]

[a] Department of Psychology, University at Buffalo, The State University of New York, Park Hall, Buffalo, NY, 14260, USA
[b] VA Center for Integrated Healthcare, VA Western NY Medical Center, Buffalo, NY, 14215, USA
[c] Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, 21287, USA
[d] Department of Psychology, University of Central Florida, Orlando, FL, 32816, USA
[e] Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, 15213, USA

ARTICLE INFO

Article history:
Received 3 May 2016
Received in revised form 10 October 2016
Accepted 12 October 2016
Available online 22 October 2016

Keywords:
Smoking
Cessation
Generalizability
Varenicline

ABSTRACT

Background: The selection criteria used in clinical trials for smoking cessation and in laboratory studies that seek to understand mechanisms responsible for treatment outcomes may limit their generalizability to one another and to the general population.
Methods: We reviewed studies on varenicline versus placebo and compared eligibility criteria and participant characteristics of clinical trials ($N=23$) and laboratory studies ($N=22$) across study type and to nationally representative survey data on adult, daily USA smokers (2014 National Health Interview Survey; 2014 National Survey on Drug Use and Health).
Results: Relative to laboratory studies, clinical trials more commonly reported excluding smokers who were unmotivated to quit and for specific medical conditions (e.g., cardiovascular disease, COPD), although both study types frequently reported excluding for general medical or psychiatric reasons. Laboratory versus clinical samples smoked less, had lower nicotine dependence, were younger, and more homogeneous with respect to smoking level and nicotine dependence. Application of common eligibility criteria to national survey data resulted in considerable elimination of the daily-smoking population for both clinical trials (≥47%) and laboratory studies (≥39%). Relative to the target population, studies in this review recruited participants who smoked considerably more and had a later smoking onset age, and were under-representative of Caucasians.
Conclusions: Results suggest that selection criteria of varenicline studies limit generalizability in meaningful ways, and differences in criteria across study type may undermine efforts at translational research. Recommendations for improvements in participant selection and reporting standards are discussed.

1. Introduction

Cigarette smoking is the leading cause of preventable death in the United States (Centers for Disease Control and Prevention, 2014). Despite availability of several pharmacological smoking cessation aids, including nicotine replacement therapy (NRT), varenicline, and bupropion, cessation rates at the population level in the United States have not increased over recent decades (Zhu et al., 2012). When results from efficacy studies fail to generalize to the population of interest, it is difficult to determine whether such failures are due to complications in properly implementing efficacious treatments, lack of efficacy in the broader population, or other factors (Luce et al., 2009). In the smoking cessation literature, there is no consensus regarding the effectiveness of pharmacotherapies assessed at the population level (see Kasza et al., 2012). Although failures to translate research findings into practice are likely due to several factors, including underutilization of available treatments (Cokkinides et al., 2005), researchers have become increasingly concerned about the generalizability of pharmacotherapy studies,

* Corresponding author.
E-mail address: stiffany@buffalo.edu (S.T. Tiffany).

http://dx.doi.org/10.1016/j.drugalcdep.2016.10.018
0376-8716/

both within (e.g., Le Strat et al., 2011) and outside (e.g., Zimmerman et al., 2005) the field of addiction.

Researchers have primarily raised concerns about generalizability of randomized-controlled trials (RCTs), which evaluate the efficacy of a medication relative to placebo and/or another efficacious treatment. These concerns arise, in part, from the extensive eligibility criteria in most RCTs, which frequently exclude individuals who have comorbid psychiatric or medical problems (e.g., Humphreys and Weisner, 2000; Le Strat et al., 2011). Limitations in generalizability of RCTs may partially account for the discouraging rates of smoking cessation observed at the population level (Le Strat et al., 2011; Luce et al., 2009). The presence of comorbid psychiatric and medical conditions, for example, may complicate the course of treatment, resulting in effectiveness rates that fall far below the original investigations of medication efficacy.

The extent to which studies on substance use disorders (SUDs) are limited in their generalizability has been assessed primarily through three methods: (1) applying common eligibility criteria to nationally representative samples and quantifying proportions of the target population that are excluded (e.g., Blanco et al., 2008; Le Strat et al., 2011), (2) applying eligibility criteria to treatment samples and quantifying proportions of the target population that are excluded (e.g., Humphreys and Weisner, 2000), and (3) comparing characteristics of research samples to treatment samples (Carroll et al., 1999; Humphreys and Weisner, 2000). These studies suggest that eligibility criteria can considerably limit generalizability, such that a large proportion of individuals with the SUD under investigation are not represented in clinical trials, with maximum exclusion rates ranging between 50% and 80% (Blanco et al., 2008; Okuda et al., 2010). Some studies have also demonstrated differences between samples of those eligible versus ineligible to participate in RCTs on age, socio-economic status, ethnicity, and comorbid psychiatric/medical and substance-use problems (Carroll et al., 1999; Humphreys and Weisner, 2000). The findings of these studies have been mixed, likely due to differences in methodology and the targeted SUD. Only one previous study has examined generalizability of smoking pharmacotherapy RCTs, (Le Strat et al., 2011), so additional research on this topic is needed.

Maximizing the effectiveness of pharmacotherapies for SUDs also requires researchers to determine how these medications achieve their desired effects. Although RCTs may include investigations of addiction-related processes (e.g., craving, withdrawal symptoms), they are primarily designed to measure efficacy via abstinence-based outcomes. Laboratory mechanistic-process studies (LMPs) are specifically designed to assess mechanisms or processes responsible for treatment outcomes. However, researchers rarely examine generalizability concerns among LMPs (c.f., Kamholz et al., 2009) despite availability of population-level data published for such purposes (Hughes, 2004; Hughes and Callas, 2010). It is important that pharmacotherapy LMPs generalize not only to the broader population of individuals with SUDs, but also to the population in which efficacy was initially demonstrated. This is a necessary requirement from a research methodology standpoint, in that mechanisms of efficacy are best investigated with samples comparable to those used to establish efficacy. Moreover, generalizable findings can produce a common knowledge base upon which greater understanding of mechanistic actions can be translated into improved clinical outcomes. To our knowledge, no previous reviews in the addiction field have comparatively examined generalizability of LMPs to RCTs in either eligibility criteria or participant characteristics (e.g., demographics, substance use history).

This review had two aims: a) to examine how eligibility criteria of LMPs and RCTs may differentially limit generalizability to smokers likely to receive medication in the "real world," and b) to assess the correspondence between participant characteristics of LMP and RCT samples. We conducted a literature review of empirical studies on varenicline, an FDA-approved partial nicotine agonist believed to aid smoking cessation by reducing cigarette craving and the reinforcing effects of smoking (Rollema et al., 2010). Varenicline is well established as a first-line treatment; with over 2.1 million prescriptions of (branded) Chantix® written annually, it is among the top 100 most prescribed medications in the USA (Brooks, 2014). Our review focused on only one pharmacotherapy to account for variability in eligibility criteria that may reflect properties of the specific medication being investigated. We limited the review to studies on randomized varenicline versus placebo, as these studies provide the most systematic means of assessing a pharmacotherapy's efficacy and mechanisms of action.[1]

To address the first aim, eligibility criteria were compared across LMPs and RCTs. Differences across study type would suggest that LMPs and RCTs were examining varenicline effects in different populations of smokers. Commonly reported eligibility criteria of LMPs and RCTs were then applied to nationally representative USA samples of daily smokers to determine proportions of the target population that would be excluded from these research studies. Examination of restrictions in representativeness was also assessed across study type. To address the second aim, participant characteristics of LMP and RCT samples were compared across study type. In addition, their characteristics were compared with nationally representative USA samples of daily smokers to determine whether these studies were using samples that markedly deviated from the population of interest.

We hypothesized that RCTs would be more restrictive on medically-related variables (Ingenito and Brewer, 2011). Based on prior research on generalizability of smoking pharmacotherapy RCTs (Le Strat et al., 2011), we hypothesized that both RCTs and LMPs would be limited in their representativeness of USA daily smokers. For participant characteristics, we sought to examine: a) whether LMP and RCT samples were, on average, generally similar to or different from the daily-smoking population on characteristics such as cigarettes smoked per day (CPD), age, etc., and b) the relative homogeneity of LMP versus RCT samples. Greater homogeneity in participant characteristics might suggest that these studies recruited a small subset of individuals on a particular characteristic (e.g., individuals who smoke 20–22 CPD), and are therefore underrepresentative of the population. We hypothesized that LMP samples would be more homogeneous than RCT samples.

2. Methods

2.1. Literature review

2.1.1. Selection of studies. A literature search conducted in the Cochrane Database of Systematic Reviews for articles on *varenicline*, and in PubMED and PsycINFO with limiters *human* and *English*, for records available online published from January 1, 2006 to May 31, 2015, yielded 849 unique records. Co-authors reviewed abstracts to select those that had empirical data, human, adult participants, sample size >1, at least some participants who were smokers, and a specific analysis conducted for smokers. Of the 849 records, 232 full articles (27%) were considered further for inclusion.

Reviewers generated definitions for RCTs using criteria from Cochrane reviews of RCTs on varenicline (Cahill et al., 2012) as a preliminary guideline and definitions for LMPs through group discussion. Reviewers dual-rated articles for inclusion based on these definitions. All studies were required to include descriptive statistics for smokers, have random assignment to and blind administration of varenicline or placebo (or randomization to order for

[1] Other active-drug conditions did not exclude a study from the review, but only participant characteristics pertaining to varenicline versus placebo were extracted.