

# Prevalence and predictors of research participant eligibility criteria in alcohol treatment outcome studies, 1970–98

Keith Humphreys, Kenneth R. Weingardt, Doyanne Horst, Asha A. Joshi & John W. Finney

Veterans Affairs and Stanford University Medical Centers, Palo Alto, California

---

*Correspondence to:*

Keith Humphreys  
Program Evaluation and Resource Center  
VAPAHS (152-MPD)  
795 Willow Road  
Menlo Park  
CA 94025  
USA  
Tel: 650 617 2746  
Fax: 650 617 2736  
E-mail: knh@stanford.edu

Submitted 22 September 2004;  
initial review completed 15 December 2004;  
final version accepted 18 April 2005

---

## ABSTRACT

**Aims** To describe the eligibility criteria (i.e. study participant inclusion and exclusion rules) employed in alcohol treatment outcome research and to identify predictors of their use.

**Design** The eligibility criteria of 683 alcohol treatment outcome studies conducted between 1970 and 1998 were coded reliably into 14 general categories. Predictors of the use of eligibility criteria were then examined.

**Findings** Patients were most often ruled ineligible for research studies because of their level of alcohol problems (39.1% of studies), comorbid psychiatric problems (37.8%), past or concurrent utilization of alcohol treatment (31.8%), co-occurring medical conditions (31.6%), and because they were deemed non-compliant and unmotivated (31.5%). The number of eligibility criteria employed in studies increased from the 1970s through the 1990s, and was positively associated with funding from the US National Institute of Alcohol Abuse and Alcoholism (NIAAA) and from the private sector, lack of an inpatient/residential treatment condition, presence of a pharmacotherapy, and use of a randomized, multiple-condition design. Principal investigators with doctoral degrees used more eligibility criteria than those with lower degrees.

**Conclusion** Participant eligibility criteria are extensively employed in alcohol treatment outcome research, and vary significantly across historical periods, funders and research designs. Researchers should report the details of subject eligibility criteria and excluded patients more fully, and, evaluate how eligibility criteria affect the cost, feasibility, and generalizability of treatment outcome research.

---

## INTRODUCTION

Alcohol treatment outcome research aspires not only to rigor but also to relevance. How well outcome studies achieve the latter of these aims can be assessed in a variety of ways, but at least one important criterion is whether research results can reasonably be generalized to everyday practice settings. Because similarity of research subjects to real-world alcohol patients may facilitate such science-to-practice linkages, it is unsettling that alcohol treatment outcome research subjects differ markedly from typical alcohol treatment seekers. For example, research participants are about 50% more likely

to be employed, more than twice as likely to be married, and three times as likely to have a college degree. Research samples also include lower proportions of women and racial minorities than do typical alcohol patient caseloads (Humphreys 2003).

Such sizable differences raise at least three questions for the field: (1) Why are alcohol treatment outcome research subjects dissimilar from typical alcohol treatment patients? (2) Do such differences matter? For example do they bias outcome findings or reduce their clinical utility? and (3) if careful evaluation proves these differences are problematic, what can be done about them? The present paper is part of a larger research program

that approaches these questions with a focus on eligibility criteria, i.e. the rules scientists employ to exclude some patients from treatment outcome studies. Eligibility criteria are a useful venue to address questions of generalizability (aka 'external validity') because researchers have more power to influence exclusion/inclusion rules than they do other factors that affect the clinical relevance of research (e.g. the reality that patients who simply refuse to enroll in research studies probably differ from those who participate).

Eligibility criteria are clearly necessary in some outcome studies. Most obviously, some patients cannot be enrolled in research for physical safety reasons, for example when a pharmacotherapy being evaluated has a documented adverse interaction with another medication a patient is taking. Ethical concerns also mandate that some individuals be excluded from research, for example alcoholic patients with Korsakoff's syndrome who cannot understand a consent form. The advantages and disadvantages of other types of eligibility criteria, for example those that are intended to make a study less costly or to facilitate data interpretation by creating a homogenous sample, have been much debated (Chalmers 1990; Fuks *et al.* 1998; Stirman *et al.* 2003). The purpose of this paper is neither to revisit nor resolve those debates. Rather, we hope to inform dialogue in this area by providing basic empirical information about how and under what conditions subject eligibility criteria have been used in alcohol treatment outcome research.

A prior study in this research program indicated that eligibility criteria may significantly affect the external validity of alcohol treatment research. Humphreys & Weisner (2000) found that most patients in a sample of 593 real-world alcohol treatment-seekers would not be eligible to enroll in an outcome study that employed common eligibility criteria (e.g. comorbid drug dependence, prior unsuccessful alcohol treatment) and that excluded patients tended to be African-American, low income and to have severe alcohol, drug, and psychiatric problems. These findings caution against assuming that outcome research is broadly generalizable and raise ethical concerns about the potential for disproportionate exclusion of disenfranchised populations from treatment research.

The above findings made worthwhile an investigation into how eligibility criteria have been employed over the modern history of alcohol treatment research. The empirical basis for this study is the exhaustive treatment research synthesis of Finney and colleagues (Moyer, Finney & Swearingen 2002; Finney, Moyer & Swearingen 2003; Swearingen, Moyer & Finney 2003). This integrative review comprised all 701 locatable alcohol treatment studies reported between 1970 and 1998 that (1) included a follow-up; (2) had at least five participants in each condition; (3) included at least some participants

who were aged 18 and over, and (4) were published in English. A 'study' was defined as a unique research project rather than each separate publication from the same project. The studies were identified through multiple, overlapping literature search procedures that covered dissertations and book chapters, as well as journal articles. Each study was coded on treatment, participant, and research design characteristics. Importantly, although it was not a primary focus of the research synthesis, the coding team transcribed all text in the study reports describing eligibility criteria, which provided the raw data for the present analysis.

This paper focused on 683 of the studies identified by Finney and colleagues that evaluated alcohol treatment (i.e. all but the 18 opportunistic brief intervention studies). Using this database, we address two questions about alcohol treatment research:

- 1 What is the nature and frequency of eligibility criteria used over the past three decades?
- 2 What are the study-specific and historical predictors of use of eligibility criteria?

## METHOD

### Definition of subject eligibility criteria

All studies exclude some patients implicitly. For example, individuals who have no contact with their families are de facto not eligible for studies of behavioral family therapy. Such cases should not be considered examples of research eligibility criteria, because they do not exclude from study participation any patient to whom the research might hope to generalize (i.e. behavioral family therapy studies are not intended to generalize to the population of patients who do not receive family therapy). True subject eligibility criteria are those rules established by a research project that limit participation by patients in the relevant sampling pool, i.e. those who would normally receive the treatment being evaluated. Thus, if a treatment study is conducted at a clinic in Oslo, Norway, its lack of enrollment of Maltans seeking culturally tailored treatment for alcohol problems would not be considered an eligibility criterion because such persons would not travel from Malta to Norway to find such alcohol treatment. However, if Norwegian patients who came to the Oslo clinic from Hamar (125 kilometers away) were ineligible for a study because the research project only had the resources to conduct local follow-up interviews, that would be an eligibility criterion, i.e. something imposed by the study that excluded from research participation some portion of the relevant real-world population.

Following the example by Fuks *et al.* (1998), we use the term 'eligibility criteria' to encompass a range of cri-

teria which scientists sometimes subdivide into 'inclusion' and 'exclusion' rules. We do this because there is no genuine logical or practical difference between, for example, an inclusion criterion for males and an exclusion criterion for females, or an inclusion criteria for employment and an exclusion criteria of unemployment. Further, on a practical note, our project team's experience is that discussing eligibility criteria quickly becomes confusing if one constantly tries to translate criteria across the inclusion/exclusion dichotomy (e.g. 'When you say that a study's exclusion criteria prevented enrollment of patients who did not score below a 9 on the AUDIT, is that the same as an inclusion criteria for patients who did score below that level . . . or would it be above that level?').

### Coding of eligibility criteria

Our system of coding eligibility criteria was developed through an iterative process over the course of several months. Initially, three study team members (KH, DH and AJ) read the verbatim eligibility criteria culled from the alcohol treatment outcome research synthesis dataset and independently attempted to arrange them into meaningful groupings. This process resulted in a long list of potential coding categories. Two raters (DH and AJ) then independently attempted to code studies using the draft categories, which resulted in a refining of the coding system as some categories were discovered to be too narrow, too broad or too confusing in practice. Through this process, the study team arrived at a system of 14 types of eligibility criteria. Subordinate categories were developed within each general type. The two independent raters then rated all the studies. The raters compared scores after each 50 studies to assess for rater drift, but these comparisons did not result in any changes in the coding system itself. Kappa for ratings of all 683 studies was 0.94, and a very high level of rater concordance was also evident in percent agreement ratings, which ranged from 95 to 100% across the 14 categories of eligibility (mean [SD] = 98.3% [1.5%]). The 14 types of eligibility criteria and their most frequently coded subordinate categories are presented in Table 1.

### Study characteristic data

To examine the correlates of the use of research participant eligibility criteria, this study employed data originally gathered by Finney and colleagues in their systematic review (Finney *et al.* 2003; Moyer *et al.* 2002; Swearingen *et al.* 2003). These data included the study's year of publication, country of origin, whether or not it included an inpatient/residential treatment condition, whether or not it included a pharmacological treatment

condition, whether or not it randomly assigned patients to experimental conditions, and whether it included multiple treatment conditions or only a single condition. Primary funding source was coded according to the following hierarchy: US National Institute on Alcohol Abuse and Alcoholism (NIAAA), US Department of Veterans Affairs (VA), other national agency, state/provincial/local government, foundation, treatment program or university, and private sector (over three-fourths of studies in this funding category were primarily funded by pharmaceutical companies, but studies supported by alcohol beverage industry funds and private gifts were also included). Finally, because the way scientists are trained seems likely to influence how they conduct research, the principal investigator's highest level of education (MD, PhD, or other) was recorded in the 533 studies where it was determinable.

## RESULTS

### Frequency of each type of eligibility criteria

Of the 683 studies, 173 (25.3%) did not mention eligibility criteria, indicating that they either had none or did not describe those they employed. Of the 510 studies that did report use of eligibility criteria, 88 (17.3% of 510) used one of the 14 types of criteria, 76 (14.9%) used two types of criteria, 81 (15.9%) used three, 91 (17.8%) used four, 59 (11.6%) used five, 47 (9.2%) used six, and 68 (13.4%) excluded potential subjects based on seven or more of the 14 types of eligibility criteria.

The prevalence of eligibility criteria varied by type (see Table 1). The most common criterion was *alcohol problems*, which was used in 39.1% ( $n = 267$ ) of the 683 outcome studies. These criteria were described in different ways (e.g. 'problematic alcohol use', 'diagnosis of alcoholism'), which is not surprising given that the research synthesis spanned a number of countries and several revisions of diagnostic criteria. In most cases, patients were ruled ineligible for research studies if their alcohol problem was not severe enough (e.g. did not meet diagnostic criteria). However, 51 studies (7.5%) excluded patients whose alcohol problems were too severe. These studies focused primarily on helping patients to achieve moderate drinking outcomes rather than abstinence.

*Psychiatric problems* emerged as the second most frequently reported type of eligibility criteria. Of the 683 projects, 258 (37.8%) did not allow enrollment of patients with psychiatric comorbidities. Some studies had specific eligibility criteria, such as the 115 that excluded patients with psychotic spectrum disorders (e.g. schizophrenia). But many studies ( $n = 95$ ) used non-specific terms to describe conditions that made patients ineligible,

**Table 1** Prevalence of categories of eligibility criteria used in 683 alcohol treatment outcome studies.

<i>Eligibility criteria</i>	<i>% of studies using</i>	<i>Most common criteria within category (n of studies)</i>
Alcohol Problems	39.1%	Must be 'alcohol dependent' ( <i>n</i> = 114) Must meet formal diagnostic criteria, e.g. for 'alcoholism' ( <i>n</i> = 70) Cannot have an alcohol problem of high severity ( <i>n</i> = 51)
Psychiatric Problems	37.8%	Cannot have a co-occurring psychotic spectrum disorder ( <i>n</i> = 115) Cannot have co-occurring 'psychiatric problems' ( <i>n</i> = 95) Cannot pose a threat to self or others or be gravely disabled ( <i>n</i> = 20) Cannot be taking psychiatric medication ( <i>n</i> = 20)
Alcohol Treatment	31.8%	Cannot be receiving concurrent services in another treatment program ( <i>n</i> = 34) Must have had prior inpatient treatment ( <i>n</i> = 27) Cannot have been treated for alcohol problems in the past ( <i>n</i> = 24)
Medical Conditions	31.6%	Cannot have chronic medical problems that might interfere with engagement in alcohol treatment ( <i>n</i> = 119) Cannot be pregnant, nursing or not using contraception ( <i>n</i> = 45) Cannot have liver disease ( <i>n</i> = 40)
Compliance/ Motivation	31.5%	Must accept mandatory attendance as a condition of treatment ( <i>n</i> = 49) Must be judged cooperative and sufficiently desirous of treatment ( <i>n</i> = 43) Must seem likely to complete treatment ( <i>n</i> = 35) Must agree to be available for follow-up as a condition of treatment ( <i>n</i> = 31) Must show willingness to follow study protocol ( <i>n</i> = 29)
Demographic	26.2%	Must be between the endpoints of an age range, e.g. 21–60 years of age ( <i>n</i> = 69) Cannot be younger than 19 ( <i>n</i> = 27) Cannot be 65 or older ( <i>n</i> = 15) Must be male ( <i>n</i> = 48)
Neurocognitive Problems	23.0%	Cannot have brain impairment, gross organicity or a history of head trauma ( <i>n</i> = 116) Cannot have neurological disorders ( <i>n</i> = 36) Cannot have unusually low IQ ( <i>n</i> = 25)
Illicit Drug Use	22.7%	Cannot be 'abusing or dependent' on illicit drugs ( <i>n</i> = 54) Cannot be using illicit drugs ( <i>n</i> = 22) Cannot have 'abused' or be 'dependent' on heroin or other illicit opiates ( <i>n</i> = 17)
Social Stability	14.9%	Must have at least one collateral contact ( <i>n</i> = 24) Must be employed ( <i>n</i> = 21) Must be married ( <i>n</i> = 18)
Distance From Treatment	10.1%	Must live within a designated distance, e.g. 30 km, from treatment facility ( <i>n</i> = 38) Must live in same city/county/state/province as treatment facility ( <i>n</i> = 28)
Residential Stability	8.6%	Must be able to provide a mailing address, telephone number or other consistent point of physical contact ( <i>n</i> = 25) Cannot be homeless or lack a fixed residence ( <i>n</i> = 21)
Education/Literacy	4.4%	Must be fluent in spoken/written English ( <i>n</i> = 24) Must be above some minimal formal educational level ( <i>n</i> = 10)
Legal Problems	3.5%	Cannot have pending legal proceedings ( <i>n</i> = 5) Cannot have felony assault history ( <i>n</i> = 4)
Financial Situation	1.3%	Must be able to pay for treatment ( <i>n</i> = 4) Must have a particular package of health insurance benefits ( <i>n</i> = 2)

such as 'psychiatric problems', 'poor psychological functioning', and the like.

The third most commonly used category of eligibility criteria (31.8% of studies,  $n = 217$ ) addressed patients past and concurrent participation in *alcohol treatment*. Prevalent examples included studies that would not enroll patients who were concurrently receiving services in another alcohol treatment program ( $n = 34$ ) and patients who had prior alcohol treatment admissions ( $n = 24$ ). Other studies in this category required that patients had received inpatient treatment prior to study enrollment ( $n = 27$ ).

Patients with *medical conditions* were also frequently ineligible for alcohol treatment studies, with 216 (31.6%) studies excluding patients on this basis. This category included 45 studies that did not allow enrollment of female patients who were pregnant, nursing or not using contraception at the time of intake.

Nearly one-third (31.5%) of the outcome studies would only enroll alcohol patients who seemed *compliant/motivated*. To be eligible for studies in this category, patients had to be unusually and visibly cooperative with the treatment research staff and protocol. This category was notable for its dearth of reference to standardized instruments to assess likely compliance or motivation, being based apparently on investigators' subjective judgments about patients (e.g. studies in this category described excluding patients for being 'difficult', 'poorly motivated', etc.).

A total of 179 (26.2%) used *demographic* criteria to limit patients' eligibility to participate, most commonly a patient's age. Gender was also sometimes used as an eligibility criterion with 48 studies only enrolling male patients.

*Neurocognitive problems* made patients ineligible for research participation in 157 studies (23% of the sample). Cognitive deficits caused by brain impairment, gross organicity, or a history of head trauma was the most common example ( $n = 116$  studies), with neurological problems (e.g. chronic headache, blackouts, peripheral neuropathy, dementia, impaired motor function, and epileptic seizures) being second ( $n = 36$ ).

Interestingly, 155 (22.7%) of the 683 studies in the alcohol treatment outcome database did not allow participation by patients who *used illicit drugs*. Most of these criteria applied to all drug use, but some were more specific (e.g. 17 studies specified heroin and other illicit opiates).

*Social stability* was employed as an eligibility criteria in 102 studies (14.9%). Broadly speaking, studies with this type of eligibility criteria prevented enrollment of socially isolated and unstable patients. Living within some specified *distance from the treatment facility* was an eligibility criterion in 10.1% of the studies ( $n = 69$ ), with distance sometimes specified as a threshold distance (e.g. 30 km) and sometimes as a geographical unit (e.g. within city or

county limits). Other aspects of patients' living situation affected study eligibility under the criterion of *residential stability* in 8.6% of the outcome studies. This type of criteria was generally aimed at preventing enrollment of transient and homeless patients. Finally, 30 studies (4.4%) required research participants to have some specified level of *education/literacy*, 24 (3.5%) excluded patients who had legal problems, and only nine studies (1.3%) used patients' *financial situation* to determine study eligibility.

### Exploratory analyses of factors predicting use of eligibility criteria

The study characteristic information was gathered by Finney and colleagues long before this study was conceived, which constrained *a priori* hypothesis testing about the sources of exclusion criteria. At the same time, it seemed reasonable to imagine that any of the study characteristics might be associated with use of eligibility criteria, so analysis of each was conducted in an explicitly exploratory manner.

Table 2 presents Kruskal-Wallis tests examining predictors of the number of the 14 eligibility criteria categories employed. This test is preferable to the more common analysis of variance procedure in cases such as this where variable distributions are skewed and heteroscedasticity is large. Decade of publication was strongly related to number of reported eligibility criteria ( $P < 0.001$ ). An accelerating increase in reported use of eligibility criteria was evident over the past three decades, with studies published in the 1990s (mean [SD] = 3.96 [2.80]) reporting almost twice as many criteria as those in the 1970s (mean [SD] = 2.06 [2.06]).

Studies conducted in France, Germany or Italy used the largest number of categories of eligibility criteria (mean[SD] = 3.79[2.55]), whereas studies conducted in England, Scotland and Ireland averaged 2.41 types of criteria (SD = 2.49), and studies in 'other' countries (typically developing nations) averaged 2.45 (SD = 2.79). But the overall test for country of origin as a predictor was just above the threshold for statistical significance ( $P = 0.098$ ).

The analysis of funding source was significant ( $P < 0.001$ ). Eligibility criteria were used most extensively in studies supported by the NIAAA (mean [SD] = 4.36 [2.73]) and by private sector funding (mean [SD] = 4.18 [2.76]). Studies not indicating funding, which, except in a few cases, may be assumed not to have had any, imposed fewer limitations on study eligibility (mean [SD] = 2.19 [2.29]), as did studies funded by treatment programs and universities (mean [SD] = 2.21 [1.91]).

Fewer eligibility criteria categories were used in studies evaluating an inpatient or residential treatment condition than in those without a 24 hour care condition

**Table 2** Kruskal-Wallis test of factors predicting number of research participant eligibility criteria categories employed in alcohol treatment outcome studies.

	<i>n</i>	<i>Number of categories of eligibility criteria mean (standard deviation)</i>	<i>chi-square value, df</i>	<i>P-level</i>
<b>Decade of Publication</b>			58.85	< 0.001
1970s	206	2.06 (2.06)	df = 2	
1980s	249	2.53 (2.28)		
1990s	228	3.96 (2.80)		
<b>Country</b>			10.70	0.098
USA	433	2.84 (2.47)	df = 6	
England, Scotland and Ireland	54	2.41 (2.49)		
Canada	65	3.25 (2.85)		
Sweden, Norway and Denmark	27	3.22 (2.55)		
Italy, German and France	33	3.79 (2.55)		
Australia and New Zealand	33	2.48 (2.27)		
Other	38	2.45 (2.79)		
<b>Funding Source</b>			65.53	< 0.001
US NIAAA	100	4.36 (2.73)	df = 7	
US Department of Veterans Affairs	43	3.42 (2.31)		
Other federal agency	112	2.88 (2.41)		
State/Provincial/Local Agency	38	3.11 (2.65)		
Foundation	22	2.86 (2.46)		
Treatment program or University	24	2.21 (1.91)		
Private sector	33	4.18 (2.76)		
Not indicated	311	2.19 (2.29)		
<b>Inpatient/Residential Treatment</b>			17.75	< 0.001
Included in study	286	2.34 (2.25)	df = 1	
No inpatient/residential condition	397	3.24(2.66)		
<b>Pharmacologic Treatment</b>			18.98	< 0.001
Included in study	108	3.84 (2.62)	df = 1	
No pharmacologic condition	575	2.68 (2.48)		
<b>Research Design I</b>			78.42	< 0.001
Randomized experiment	282	3.88 (2.57)	df = 1	
Other design	401	2.15 (2.26)		
<b>Research Design II</b>			43.15	< 0.001
Multiple Conditions	386	3.40(2.61)	df = 1	
Single Condition	297	2.17(2.25)		
<b>Highest Education of PI</b>			9.13	0.010
MD and MD/PhD	157	3.04 (2.50)	df = 2	
PhD	329	2.83 (2.55)		
Other	47	1.83 (2.03)		

( $P < 0.001$ ), whereas more types of eligibility criteria were used in studies that included a pharmacological treatment than in non-pharmacologic studies ( $P < 0.001$ ). Research designs that included random assignment to treatment conditions ( $P < 0.001$ ) and multiple treatment conditions ( $P < 0.001$ ) were also strongly associated with a greater number of eligibility criteria.

Finally, principal investigator educational level was significantly associated ( $P = 0.010$ ) with reported use of eligibility criteria. MDs and PhDs both averaged around three types of eligibility criteria in their studies, but those with other degrees (typically Master's degrees) averaged only 1.83 (SD = 2.03) criteria.

To determine whether predictive patterns varied across individual categories of eligibility criteria, we conducted chi-squares to assess the association of each of the above predictors with each of the six most prevalent criteria (results not shown). These individual results showed the same pattern as the above results for total number of criteria, making it seem reasonable to predict total number of eligibility criteria as a general dimension in a multivariate model.

To this end, we constructed an ordinary least squares (OLS) regression model predicting the total number of categories of eligibility criteria used by each study. All significant univariate predictors as described above were

candidates for inclusion in the model. However, the variable indicating whether the study included multiple treatment conditions (Research Design II in Table 2) was not included because it was collinear ( $r = 0.61$ ) with the variable reflecting whether the study was a randomized experiment. Similarly, the presence of a pharmacotherapy was not included because it was collinear ( $r = 0.36$ ) with private sector funding. The absolute values of the correlations among the other predictors were all lower than 0.3, reflecting sufficient independence to be included in the regression model. To maximize statistical power, principal investigator's level of education was not included because it was not reported for 160 studies.

The included predictors were whether the study: (1) was published in the 1970s, 1980s, or 1990s; (2) was supported by the NIAAA; (3) had private sector funding; (4) was a randomized experiment and (5) included an inpatient/residential treatment condition. All but the last of these five predictors remained significant in the multivariate model (Table 3). For the four dichotomous predictors, the unstandardized *b*-weights reflect the change in number of eligibility criteria types predicted by the study characteristic being present rather than absent, holding the other predictors constant. The presence of random assignment (1.39 more criteria types) and funding from the NIAAA (1.22 more criteria types) were the strongest dichotomous predictors of greater use of eligibility criteria. The *b*-weight for the trichotomous historical period predictor variable indicates that with each decade from the 1970s to the 1990s, treatment outcome research studies increased their average number of eligibility criteria by 0.77 over the preceding decade.

## DISCUSSION

All conclusions from this study must be interpreted in light of the data source, namely those study characteristics coded by Finney and colleagues, which in turn were dependent on what was presented in the original scientific reports. Some study variables (e.g. urbanicity of

study location) were not coded by Finney and colleagues' original project, so their potential influence could not be assessed in the present analysis. Even for those variables that were coded, many important aspects of research design are inadequately described in published alcohol treatment research, and this problem extends to eligibility criteria (Wilk, Jensen & Havinghurst 1997; Moncrieff & Drummond 1998; Moyer *et al.* 2002). It follows that our finding that about three-fourths of outcome studies use at least some eligibility criteria is a conservative estimate, because some of the remaining fourth probably employed eligibility criteria but failed to report them.

This situation leads us naturally to the recommendation that all studies clearly describe their eligibility criteria and, just as importantly, report on the number of patients excluded from enrollment on the basis of their criteria. Moncrieff & Drummond (1998) found that of the few alcohol studies that did such reporting, the average study excluded half of all patients, with a high of 92%! If reporting such information were mandatory, as is increasingly being advocated in the medical research community (see Moher *et al.* 2001), one wonders whether many conclusions drawn in discussion sections would have to be tempered, and, in parallel fashion, whether some 'bench-to-bedside' technology transfer efforts would become more cautious in assuming that what is learned in clinical trials can be safely applied in real-world clinical practice.

In addition to being better described, eligibility criteria should be implemented using objective measures. If eligibility criteria are not explicit and standardized, independent investigators cannot scientifically replicate outcome findings. The most worrisome of the subjective criteria identified in this project are those that excluded patients whom the investigators judged 'difficult', 'unmotivated', 'unlikely to benefit from treatment', etc. In some outcome studies, the evaluator has a strong personal stake in the treatment, which may be a matter of financial reward (e.g. the desire to get a new medication to market) or enhanced reputation (i.e. when the author is the creator of a well-known therapy being evaluated, or the director of the treatment program being studied). In such situa-

**Table 3** OLS regression model predicting number of research participant eligibility criteria employed in 683 alcohol treatment outcome studies.

Predictor	<i>b</i> -weight	SE	<i>t</i>	<i>P</i>
Constant	0.60	0.27	2.23	0.026
Decade of publication	0.77	0.11	6.93	0.000
US NIAAA funding	1.22	0.25	4.87	0.000
Private sector funding	0.82	0.41	1.99	0.047
Randomized experimental design	1.39	0.19	7.48	0.000
Inpatient/Residential treatment condition	-0.21	0.18	-1.13	0.259

Model  $F = 38.31$ ,  $P < 0.001$ . Decade of publication coded 1 = 1970s, 2 = 1980s, 3 = 1990s. All other variables coded 0 = No, 1 = Yes.

tions, vague eligibility criteria subjectively applied can be a temptation to prefigure a study's outcome through the selection of 'easy cases'.

As a final comment on the quality of scientific reporting, degree of design description enters as a confound in one of the largest associations identified here: the steadily rising number of outcome study eligibility criteria types employed over the past three decades. Did eligibility criteria really become more prevalent from the 1970s to the 1990s, or did they just become more frequently acknowledged in research articles? Quality of scientific reporting has risen somewhat over time; for example 42.1% of alcohol outcome studies conducted in the 1990s reported the number of patients approached to participate, compared to 35.3% in the 1970s (Moyer *et al.* 2002). But the association identified here between decade and number of eligibility criteria reported is too large to be fully explained by the relatively modest improvements in quality of scientific reporting.

Funding by the NIAAA and by private sources both predicted greater use of eligibility criteria. In the case of NIAAA funding, which was significant in both the univariate and multivariate analyses, this may reflect that for most of the study period, NIAAA did not have a 'health services research' focus, i.e. a legitimated program of 'effectiveness' studies that emphasized evaluations under real-world conditions. Rather, tightly controlled 'efficacy' trials were the norm, and such studies usually excluded many patients. Privately funded studies (as well as publicly funded pharmacotherapy studies) may use more eligibility criteria because they often involve medications for which there are medical and psychiatric contraindications. At least for pharmaceutical companies, an added explanation may be regulations for how medications are approved. During the study period, the US Food and Drug Administration approved medications based on significant effects found in a study whether the treatment group was representative or not, so there were no penalties attendant to excluding from pharmaceutical studies all but the most compliant and healthy patients.

Studies that examined inpatient and residential treatments had fewer eligibility criteria, perhaps because project teams and institutional review boards are more comfortable enrolling patients with serious comorbidities in a study within a 24 hour treatment setting than in, say, an outpatient clinic that has only a few hours of contact with patients each week. A non-competing explanation is that the high cost of inpatient treatment excludes most disadvantaged and troubled patients (Monahan & Finney 1996), making further exclusions by an evaluator superfluous.

The greater use of eligibility criteria in randomized trials and multiple condition studies may help explain a

problem in technology transfer, which was illustrated at a recent conference for practitioners at which Project MATCH (1997) results were presented. After listening to a detailed presentation of the research design and results of this controlled clinical trial (which had extensive eligibility criteria), an audience member asked the presenter in an exasperated tone 'What does any of this have to do with *treatment?*' This challenge met with broad approval among the audience (Dr Dennis Donovan, personal communication, 11 August 2003). Random assignment per se has been shown to exert only a modest impact on addiction treatment outcomes (McKay *et al.* 1995, 1998); perhaps extensive eligibility criteria are the research design feature that better accounts for many clinicians' doubts that clinical trials have something 'to do with treatment.' Whether outcome studies' exclusion of patients typically seen in everyday practice is a basis of clinical skepticism is a question we intend to address as we continue with this research program.

Finally, it remains to be determined whether extensive use of eligibility criteria makes an important difference (i.e. is clinical skepticism of outcome research appropriate or misplaced?). We feel fairly confident that at least some eligibility criteria have little effect on generalizability. For example, many studies examined here required a formal diagnosis of alcohol dependence/alcoholism for study eligibility. This criterion may have excluded some patients with subdiagnostic alcohol problems, which could somewhat reduce clinical relevance, but our inspection of these studies showed that most were conducted in large, urban medical facilities with caseloads including patients who had no genuine alcohol problems (e.g. homeless patients seeking a bed and meals, drug dependent patients who do not drink alcohol, psychiatric patients seeking care and safety). On balance, such an eligibility criteria would, if anything, enhance clinical relevance by enrolling only those individuals who ought to receive alcohol treatment in the real world.

However, most of the eligibility criteria identified here raise concerns that are not as readily resolved. Obviously, if most patients are ineligible for enrollment, research recruitment periods must be extended, which can pose fiscal and practical problems for a study. But the more pressing question is whether eligibility criteria significantly change the outcome results a study obtains. Judged at a glance, most of the prevalent eligibility criteria in alcohol treatment studies are likely to exclude from research participation more poor prognosis than good prognosis patients. Some studies suggest that members of disenfranchised groups are more likely to be excluded from alcohol research under commonly applied eligibility criteria (Humphreys & Weisner 2000). However, because of the complex interplay of patient problems, life contexts, treatment services received, likelihood of follow-up, and

measurement issues, it is not clear what effect excluding more severely troubled patients would have on outcome results. In our view, that remains the central question for future studies in this area.

### Acknowledgements

This study was funded by NIAAA grants R-01 AA13315 and R-01 AA08689, and by the Department of Veterans Affairs Health Services Research and Development Service and Mental Health Strategic Healthcare Group. We thank Anne Moyer and Dennis Donovan for their assistance.

### References

- Chalmers, T. C. (1990) Ethical implications of rejecting patients for clinical trials. *JAMA*, **263**, 865.
- Finney, J. W., Moyer, A. & Swearingen, C. (2003) Outcome variables and their assessment in alcohol treatment studies: 1968–98. *Alcoholism: Clinical and Experimental Research*, **27**, 1671–1679.
- Fuks, A., Weijer, C., Freedman, B., Shapiro, S., Skrutkowska, M. & Riaz, A. (1998) A study in contrast: Eligibility criteria in a twenty-year sample of NSABP and POG clinical trials. *Journal of Clinical Epidemiology*, **51**, 69–79.
- Humphreys, K. (2003) Do participants in alcoholism treatment outcome studies resemble patients seen in everyday practice? *Psychiatric Services*, **54**, 1576.
- Humphreys, K. & Weisner, C. (2000) Use of exclusion criteria in selecting research subjects and its effect on the generalizability of alcohol treatment outcome studies. *American Journal of Psychiatry*, **157**, 588–594.
- McKay, J. R., Alterman, A. I., McLellan, A. T., Broadman, C., Mulvaney, F. & O'Brien, C. P. (1998) The effect of random versus non-random assignment in the evaluation of treatment for cocaine abusers. *Journal of Consulting and Clinical Psychology*, **66**, 697–701.
- McKay, J. R., Alterman, A. I., McLellan, A. T., Snider, E. C. & O'Brien, C. P. (1995) Effect of random versus non-random assignment in a comparison of inpatient and day hospital rehabilitation for male alcoholics. *Journal of Consulting and Clinical Psychology*, **63**, 70–78.
- Moher, D., Schulz, K. F. & Altman, D. (2001) The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*, **285**, 1987–1991.
- Monahan, S. C. & Finney, J. W. (1996) Explaining abstinence rates following treatment for alcohol abuse: a quantitative synthesis of patient, research design, and treatment effects. *Addiction*, **91**, 787–805.
- Moncrieff, J. & Drummond, D. C. (1998) The quality of alcohol treatment research: an examination of influential controlled trials and development of a quality rating system. *Addiction*, **93**, 811–823.
- Moyer, A., Finney, J. W. & Swearingen, C. E. (2002) Methodological characteristics and quality of alcohol treatment outcome studies, 1970–98: An expanded evaluation. *Addiction*, **97**, 253–263.
- Project MATCH, Research Group (1997) Matching alcoholism treatments to client heterogeneity: Project MATCH post-treatment drinking outcomes. *Journal of Studies on Alcohol*, **58**, 7–29.
- Stirman, S. W., DeRubeis, R. J., Crits-Cristoph, P. & Brody, P. E. (2003) Are samples in randomized controlled trials of psychotherapy representative of community outpatients? A new methodology and initial findings. *Journal of Consulting and Clinical Psychology*, **71**, 963–972.
- Swearingen, C. E. A. & Moyer & Finney, J. W. (2003) Alcoholism treatment outcome studies, 1970–98: An expanded look at the nature of research. *Addictive Behaviors*, **28**, 415–436.
- Wilk, A. I., Jensen, N. M. & Havighurst, T. C. (1997) Meta-analysis of randomized control trials addressing brief interventions in heavy alcohol drinkers. *Journal of General Internal Medicine*, **12**, 274–283.