Avens Publishing Group
J Addiction Prevention
November 2013 Vol.:1, Issue:3

Open Access **Research Article**

Journal of
Addiction &
Prevention

Flexibility in Design, Outcomes and Analysis in "Evidence-Based" Drug Prevention Research: The Case of the Midwestern Prevention Project

Dennis M. Gorman*
Department of Epidemiology and Biostatistics, School of Rural Public Health, Texas A&M University, TAMU 1266, College Station, Texas 77843-1266, USA

***Address for Correspondence**
Dennis M. Gorman, Ph.D., Department of Epidemiology and Biostatistics, School of Rural Public Health, Texas A&M University, TAMU 1266, College Station, Texas 77843-1266, USA, Tel: 979-458-8059; Fax: 979-458-1877; E-mail: gorman@srph.tamhsc.edu

Submission: 24 October 2013
Accepted: 20 November 2013
Published: 22 November 2013

Keywords: Drug Prevention Programs; Alcohol; Cigarettes; Marijuana; Analytic Flexibility; Community-based Prevention Programs; Evidence-based Practice

Abstract

Flexibility in study designs, definitions, outcomes and analytic models increases the chance that the results reported in an empirical study are untrue. Such flexibility in methodological and analytic practices has been observed in evaluations of a number of drug prevention programs that appear on lists of evidenced-based programs. The current paper examines the evidence base pertaining to one of the most well-established drug prevention programs, namely the Midwestern Prevention Project (MPP; also known as Project STAR). Specifically, it examines the results reported from the quasi-experimental evaluation of the MPP that was conducted in Kansas City and the experimental replication conducted in Indianapolis, Indiana, and the data analysis practices used in producing these findings. It shows there is considerable analytical flexibility evident in the published accounts from these two evaluations studies, notably in the samples used in the data analyses and the manner in which outcome variables were measured. The implications of this for the MPP's status as an evidenced-based drug prevention program are discussed, along with the means by which flexibility in analytic procedures in drug prevention research could be reduced. The seemingly widespread existence of such analytic flexibility within the field suggests that confirmation bias is part of the general culture of drug prevention research. This impedes the development of a sound scientific base within the discipline and might even lead it to degenerate into a pseudoscience.

Introduction

Ioannidis [1] recently noted that the greater the flexibility in study designs, definitions, outcomes and analytic models, the less likely it is that research findings are true. Elaborating further, he observed that flexibility in the case of analytic methods allows "experimentation" in data analysis and the option to report only the "best" results. In the case of outcomes, flexibility is greatest when these can take many different forms and can be measured in many different ways (e.g., scales that can be dichotomized at different points). This gives the researcher more freedom to choose the measure that produces the "best" result.

A number of case studies have observed the presence of such flexibility in the methodological and analytic practices employed in evaluations of drug prevention programs [2-7]. Summarizing this literature, Holder [8] highlighted the frequent use of 1-tailed tests of statistical significance, *post hoc* outcome variable selection, obfuscation of selection bias, and exaggeration of statistically significant effects that are of little or no clinical significance. He also noted that critical debate over the use of such practices and the interpretation of study findings was largely absent from the drug prevention field and that this was likely to impede the development of a sound scientific base to the discipline.

The current paper examines in detail the evidence base pertaining to one of the most well-established drug prevention programs, namely the Midwestern Prevention Project (MPP; also known as Project STAR). The MPP developer recently described it as an "evidence-based program for prevention of both drug use and violence" [9], noting that it had been selected for inclusion on two high profile lists of model programs (the University of Colorado's *Blueprints for Violence Prevention* [10,11] and the Substance Abuse and Mental Health Service Administration's *National Registry of Evidence-based Programs and Practices* [12]). In addition, the MPP has also been included in the National Institute on Drug Abuse's *Research-based Guide* [13] and the Promising Programs Network's *Programs that Work* [14], and was listed as a model risk prevention program by the US Department of Health and Human Services in its Surgeon General's report on youth violence [15]. The fundamental contention of the present paper is that the MPP would not be considered an evidence-based or model drug prevention program were it not for the use of the type of flexible analytic practices described by Ioannidis [1].

The evidence pertaining to the MPP comes from two National Institute on Drug Abuse (NIDA) funded evaluation studies that began in the mid-1980s [16]. The first evaluation, which used a quasi-experimental design, was conducted in Kansas City and started in 1984. The second, which used an experimental design, was conducted in Indianapolis, Indiana and commenced in 1987. The Kansas City quasi-experiment has been much more extensively reported-on in the peer-review literature than the Indianapolis experiment.

The Kansas City Quasi-Experimental Evaluation of the Midwestern Prevention Project

The MPP was designed as a multicomponent community-based program with sequential introduction of five components: a mass media campaign, a school-based curriculum, a parent program, community organization involving training of city leaders in the planning and implementation of prevention efforts, and a health policy component designed to initiate change in local ordinances regulating the availability of alcohol and tobacco products [10,16,17]. The school-based component involved 10 to 13 skills training

ISSN: 2330-2178

sessions taught in 6th or 7th grade, with a five-session booster program the following year. The training was dedicated to the development of drug-specific resistance skills, and addressed the psychosocial consequences of drug use, beliefs about drug use prevalence, recognition of adult, media and community influences supportive of drug use, peer and environmental pressures, assertiveness training and problem-solving in difficult situations. The parent component was orientated around ten homework assignments which involved interviewing family members about techniques for avoiding alcohol and drug use and family methods used to counteract media and community influences. The mass media component included 16 television, 10 radio and 30 print media events, but these were present in both the intervention and comparison conditions. These program activities were implemented over the course of a 17-month period [16].

The baseline sample of the Kansas City study comprised some 5,000 students, of whom 1,607 were tracked individually over the course of the study (this figure is given as 1,606 in some publications from the study) and the remainder tracked by grade cohort (including new intake students who may not have received the intervention). The samples were assessed at baseline (September 1984) when subjects were in 6th or 7th grade (aged 12 or 13 years), and at least some of the subjects have been followed-up on 15 subsequent occasions into "early adulthood" (27 through 34 years) [17]. Despite this wealth of data collection, there have been just two periods of time in which findings from the MPP Kansas City evaluation have been published, first between 1989 and 1991 (focused on either the panel sample or grade cohort) and then, 18 years later, in 2009 (focused on the panel sample).

Effects on Cigarette, Alcohol and Marijuana Use at the 1-, 2- and 3-Year Follow-ups

The foci of the initial papers published from the Kansas City quasi-experimental evaluation of the MPP were cigarettes, alcohol and marijuana use at the 1-year follow-up [16,18-20], the 2-year follow-up [21-23] and the 3-year follow-up [24]. Tables 1 through 3 summarize the findings from these publications. Each table shows which sample was used in the data analysis reported in the paper (panel or combined) and the specific outcome for each type of drug reported (i.e., past-week use, past-month use, and, in the case of cigarettes, "other").

With the exception of Dwyer et al. [20], all of the publications state that they use 1-tailed tests of statistical significance. Use of 1-tailed tests of statistical significance in evaluating drug prevention programs doubles the threshold for rejecting the null hypothesis of no effect and involves making the claim that a program cannot produce iatrogenic effects [8,25]. However, as noted by Valente et al. [26], it is largely impossible to entirely rule out the occurrence of the latter, and iatrogenic effects have indeed been reported in evaluations of a number of prominent drug prevention programs that appear on evidence-based lists [e.g., 26-28]. Since the focus of the present paper is on how analytic flexibility can increase the chance of generating positive results it seemed reasonable to distinguish those findings reported in the MPP Kansas City evaluation that would have been statistically significant had the analyses employed the traditional 2-tailed test of statistical significance from those that would not be statistically significant using such a test (i.e., those that are only statistically significant using the more lenient 1-tailed test). In the case of six of the eight publications from the evaluation [16,18,19,21-23] this meant distinguished those results reported as statistically significant at $p < 0.01$ (using a 1-tailed test) from those results statistically significant at $p < 0.05$ (using a 1-tailed test). The former was considered a statistically significant positive effect and denoted by a "✓" in tables 1 to 3, whereas the latter was considered a questionable statistically significant effect and denoted by a "?" in the tables. Where no statistically significant effect was reported (including p values of 0.10 using 1-tailed tests, as in two of the publications [22,23]), a "×" appears in the tables. With regard to the remaining two publications, Johnson et al. [24] reported exact p values, one of which (for cigarette use) would have been statistically significant using a 2-tailed test of statistical significance, one of which (for alcohol use) would not, and one of which (for marijuana use) was of questionable statistical significance. Accordingly, a "✓", a "×", and a "?" appear in each of the respective tables. Finally, in the case of Dwyer et al. [20], no effect on alcohol use was reported (hence "×" appears in Table 2), along with a reduction in the prevalence of cigarette users and mixed evidence of an effect on marijuana use (each of which is interpreted as a positive effect in Table 1 and Table 3, respectively).

As shown in Table 1, each of the eight publications published between 1989 and 1991 contained data pertaining to cigarette use, with seven of these presenting data from the 1-year follow-up, three data from the 2-year follow-up, and one data from the 3-year follow-up. Four of the five studies that reported findings from the combined sample at the 1-year follow-up found a statistically significant effect on cigarette use during the previous month [16,19,21,22]. The effects on use during the previous week and the "other" outcome variables were less impressive as it is uncertain whether most of these would be statistically significant had 2-tailed tests been employed in the statistical analysis. Just one study [22] reported results for cigarette use at the 2-year follow-up from the combined sample, and none of the three findings presented would have been statistically significant using 2-tailed tests. Two publications reported results from the panel sample at the 1-year follow-up [20,23] and two at the 2-year follow-up [21,23]. The majority of these results were statistically significant, as was the single result reported from the 3-year follow-up [24].

Table 2 shows the results for alcohol use from the six studies that reported on this outcome. Most of the results pertained to the combined sample at the 1-year follow-up: three of these results were statistically significant, while four were of questionable significance [16,18-21]. Results from the panel sample are sparse: Dwyer et al. [20] reported results for past-week use at the 1-year follow, Pentz et al. [21] reported results for past-month use at the 2-year follow-up, and Johnson et al. [24] reported results for past-month use at the 3-year follow-up. Only the result at the 2-year follow-up was statistically significant.

Table 3 shows the results from the six publications that reported effects on marijuana use. Of the five that reported results from the 1-year follow, just one found a statistically significant effect on past-month use in the combined sample [19] and one a statistically significant effect on past-week use in the panel sample [20]. The results reported in the other three studies for the 1-year follow-up all pertain to past-week and past-month use among the combined sample and each of them is of questionable statistical significance (i.e., it is unclear whether they would be statistically significant had 2-tailed tests been used in the data analysis) [16,18,21]. One study reported results on past-month use at the 2-year follow-up [21] and one at the 3-year follow-up [24] using data from the panel sample

ISSN: 2330-2178

Table 1: Statistically Significant Program Effects on Cigarettes Use Reported in Publications from the First Three Years of the MPP Evaluation[a]

	Sample used	Year 1[d]			Year 2			Year 3		
	in Analysis[b]	Other[c]	Week	Month	Other[c]	Week	Month	Other	Week	Month
Pentz et al. [16][e]	Panel	---	---	---						
	Combined	--	?	✓						
Pentz et al. [18][f]	Panel	---	---	---						
	Combined	--	?	?						
MacKinnon et al. [19][g]	Panel	---	---	---						
	Combined	---	---	✓						
Dwyer et al. [20][h]	Panel	---	✓	---						
	Combined	---	---	---						
Pentz et al. [21][i]	Panel	---	---	---	---	---	✓			
	Combined	---	✓	✓	---	---	---			
Pentz et al. [22][j]	Panel	---	---	---	---	---	---			
	Combined	?	?	✓	×	×	?			
Pentz et al. [23][k]	Panel	✓	?	✓	?	✓	✓			
	Combined	---	---	---	---	---	---			
Johnson et al. [24][l]	Panel							---	---	✓
	Combined							---	---	---

Note: [a] --- = data analysis not reported; ✓ = statistically significant positive effect reported (i.e., *p* values < 0.01 using 1-tailed tests); ? = questionable statistically significant effect (i.e., *p* values of <0.05 using 1-tailed tests); × = no statistically significant effect reported.

[b] Combined = approximately 5,000 tracked individually and by grade cohort; Panel = 1607 students tracked individually.

[c] All of the studies report "past week" and/or "past month" use as an outcome variable. Two studies report one additional outcome measure: "past day" use [22] and "current" use [23]. These outcomes are listed under the column "other". In addition, these studies, along with Pentz et al. [18], report data pertaining to "lifetime use". This is not included in the table, as it is primarily an indicator of "background drug use experience prior to intervention" not as a dependent variable measuring program effects [18, p. 271].

[d] Two papers also contain data pertaining to a 6-month follow-up. Pentz et al. [21] contains a figure that includes 6-month follow-up data, but presents no statistical analysis of these [21, figure 1 and table 1]. Pentz et al. [23] contains statistical analysis of data from the 6-month follow-up (with results similar to those of the 1-year follow-up).

[e] Data from adjusted analysis reported in Pentz et al. [16, table 4].

[f] Data from Pentz et al. [18, table 3] which reports results for 12 schools with a high level of program implementation and 12 schools with a low level of implementation. The former results are reported in the above table as these were more favorable to the program.

[g] Data from MacKinnon et al. [19, p. 168, column 2].

[h] Dwyer et al. [20] used a number of different models in their analysis and concluded that these showed a reduction in the prevalence of cigarette users in the MPP schools compared to the controls [20, p. 781]. Hence "✓" appears in the table.

[i] Data for the 1-year follow-up are from Pentz et al. [21, table 1]. The analysis of the 2-year follow-up data does not present exact cut-offs for *p* values (e.g., *p* <0.05). Rather it is stated that the values for cigarette use, marijuana use and alcohol use each fall between <0.05 and 0.001 (using 1-tailed tests) and a figure is presented that shows prevalence of use in the last month [21, figure 1]. One cannot determine from this if the difference reported for cigarette use would be statistically significant at *p* <0.05 (using a 2-tailed test). However, since the effect at 1-year was said to be "maintained" [21, pp.220-221], a "✓" appears in the above table.

[j] Data reported are from the analysis of longitudinal program effects using endpoint schools [22, table 2, column 9 and table 3, column 9].

[k] Data pertaining to "past week" use and "past month" use are from the adjusted analysis reported in Pentz et al. [23, table 4]. Data pertaining to current use are from Pentz et al. [23, p. 721).

[l] Data from Johnson et al. [24, table 1].

[21,24]. Again, it is uncertain if these findings would be statistically significant had 2-tailed tests been employed in the analysis and hence they are designated as of questionable statistical significance in the table.

Other than the findings pertaining to use of cigarettes in the previous month, the results presented in tables 1 through 3 can hardly be said to present a compelling case that the MPP prevents drug use among adolescents. However, of most relevance to the current paper's analysis of analytic flexibility is that the focus of the Kansas City evaluation narrowed after the first-year follow-up both in terms of the subjects included in the analysis and the outcomes reported. Most notably, no results pertaining to alcohol and marijuana use were reported for the combined sample at the second or third year follow-up and the results that were reported for the panel sample were all derived from a very simple dichotomous measure of use during the previous month. Interestingly, two of the three findings reported by Pentz et al. [22] for cigarette use at the 2-year follow-up using the combined sample were not statistically significant even with the use of 1-tailed tests of statistical significance, and the focus of the evaluation for this outcome variable also shifted to the panel sample at this point. Thus, by the 3-year follow-up the data reported for all three drugs were limited to just the panel sample and to the simple

ISSN: 2330-2178

Table 2: Statistically Significant Program Effects on Alcohol Use Reported in Publications from the First Three Years of the MPP Evaluation[a]

	Sample used	Year 1[c, d]		Year 2		Year 3	
	In Analysis[b]	Week	Month	Week	Month	Week	Month
Pentz et al. [16][e]	Panel	---	---				
	Combined	?	✓				
Pentz et al. [18][f]	Panel	---	---				
	Combined	?	?				
MacKinnon et al. [19][g]	Panel	---	---				
	Combined	---	?				
Dwyer et al. [20][h]	Panel	×	---				
	Combined	---	---				
Pentz et al. [21][i]	Panel	---	---	---	✓		
	Combined	✓	✓	---	---		
Johnson et al. [24][j]	Panel					---	×
	Combined					---	---

Note: [a] --- = data analysis not reported; ✓ = statistically significant positive effect reported (i.e., *p* values < 0.01 using 1-tailed tests); ? = questionable statistically significant effect (i.e., *p* values of <0.05 using 1-tailed tests); × = no statistically significant effect reported.

[b] Combined = approximately 5,000 tracked individually and by grade cohort; Panel = 1607 students tracked individually.

[c] Only two columns of outcome variables are shown in the table for each follow-up period, since neither of the reports that included an additional outcome measure [22,23] reported any data pertaining to alcohol use. Pentz et al. [18] reported results pertaining to "lifetime use" but this is not included in the table as it is primarily an indicator of "background drug use experience prior to intervention" not as a dependent variable measuring program effects [18, p. 271].

[d] Two papers also contain data pertaining to a 6-month follow-up. Pentz et al. [21] contains a figure that includes 6-month follow-up data, but presents no statistical analysis of these [21, figure 1 and table 1]. Pentz et al. [23] contains statistical analysis of data from the 6-month follow-up (with results similar to those of the 1-year follow-up).

[e] Data from adjusted analysis reported in Pentz et al. [16, table 4].

[f] Data from Pentz et al. [18, table 3] which reports results for 12 schools with a high level of program implementation and 12 schools with a low level of implementation. The former results are reported in the above table as these were more favorable to the program.

[g] Data from MacKinnon et al. [19, p. 168, column 2].

[h] Dwyer et al. [20] used a number of different models in their analysis and concluded that these showed "no evidence of an effect on alcohol use" [20, p. 781]. Hence "×" appears in the table.

[i] Data for the 1-year follow-up are from Pentz et al. [21, table 1]. The analysis of the 2-year follow-up data does not present exact cut-offs for *p* values (e.g., *p* <0.05). Rather it is stated that the values for cigarette use, marijuana use and alcohol use each fall between <0.05 and 0.001 (using 1-tailed tests) and a figure is presented that shows prevalence of use in the last month [21, figure 1]. One cannot determine from this if the difference reported for alcohol use would be statistically significant at *p* <0.05 (using a 2-tailed test). However, since the effect at 1-year was said to be "strengthened slightly" [21, p.221], a "✓" appears in the table.

[j] Data from Johnson et al. [24, table 1].

dichotomous measure of use in the previous month. This narrowing down of drug use outcomes to just one fairly crude measure is curious since subjects were in 9th and 10th grade by this time and one might reasonably have expected some results pertaining to more regular consumption and/or heavy use of the three drugs to be reported. Indeed, since a number of subjects were using drugs by this time (34% drinking alcohol during the previous month, 25% smoking cigarettes, and 12% smoking marijuana) [24], regular and heavy use of these substances were likely occurring among the sample.

Drug Use and Other Outcomes at Later Follow-ups

Table 4 summarizes the follow-up assessments and objectives of the three recent publications that have emerged from the MPP Kansas City quasi-experiment [9,17,29]. Each of these publications analyzed data from random samples of either 1,002 or 541 subjects drawn from the 1,607 individuals in the original panel sample. One thing that immediately stands out from the table is that there have been 13 follow-up assessments of these subjects since the 3-year follow-up in 9th or 10th grade reported in Johnson et al. [24]. They were followed up on two more occasions in high school (10th/11th grade and 11th/12th grade), five times during "emerging adulthood" (19 through 26 years), and four times during "early adulthood" (26 through 30-34 years) [17]. And yet just three publications have been produced from what one can only assume to be a wealth of data. In addition, none of these three publications contains an account of the effects of the MPP on the three drugs that were the primary outcomes of the evaluation, namely cigarettes, alcohol and marijuana. Rather, what have been reported in these publications are the effects of involvement in the program on very specific outcomes assessed at various points in time during the more than 20 years of data collection. Riggs et al. [29], for example, report the effects of the MPP on amphetamine use using data from 13 of the assessment points. Growth curve analysis showed delayed initiation of amphetamine use among those who participated in the MPP compared to the control subjects and a continued reduction in the growth in amphetamine use beginning in high school (assessments 3-6) and continuing into emerging and early adulthood (assessments 7-14).

In an even more narrowly focus analysis, Riggs and Pentz [9] examined whether the MPP's effects on marijuana use during high school (assessments 4 to 6) mediated its influence on utilization of mental health services in adulthood among 961 subjects who provided data at either assessments point 13 or 14. Structural equation

ISSN: 2330-2178

Table 3: Statistically Significant Program Effects on Marijuana Use Reported in Publications from the First Three Years of the MPP Evaluation[a]

	Sample used	Year 1[c,d]		Year 2		Year 3	
	In Analysis[b]	Week	Month	Week	Month	Week	Month
Pentz et al. [16][e]	Panel	---	---				
	Combined	?	?				
Pentz et al. [18][f]	Panel	---	---				
	Combined	?	?				
MacKinnon et al. [19][g]	Panel	---	---				
	Combined	---	✓				
Dwyer et al. [20][h]	Panel	✓	---				
	Combined	---	---				
Pentz et al. [21][i]	Panel	---	---	---	?		
	Combined	?	?	---	---		
Johnson et al. [24][j]	Panel					---	?
	Combined					---	---

Note: [a] --- = data analysis not reported; ✓ = statistically significant positive effect reported (i.e., *p* values < 0.01 using 1-tailed tests); ? = questionable statistically significant effect (i.e., *p* values of <0.05 using 1-tailed tests); × = no statistically significant effect reported.

[b] Combined = approximately 5,000 tracked individually and by grade cohort; Panel = 1607 students tracked individually.

[c] Only two columns of outcome variables are shown in the table for each follow-up period, since neither of the reports that included an additional outcome measure [22,23] reported any data pertaining to marijuana use. Pentz et al. [18] reported results pertaining to "lifetime use" but this is not included in the table as it is primarily an indicator of "background drug use experience prior to intervention" not as a dependent variable measuring program effects [18, p. 271].

[d] Two papers also contain data pertaining to a 6-month follow-up. Pentz et al. [21] contains a figure that includes 6-month follow-up data, but presents no statistical analysis of these [21, figure 1 and table 1]. Pentz et al. [23] contains statistical analysis of data from the 6-month follow-up (with results similar to those of the 1-year follow-up).

[e] Data from adjusted analysis reported in Pentz et al. [16, table 4].

[f] Data from Pentz et al. [18, table 3] which reports results for 12 schools with a high level of program implementation and 12 schools with a low level of implementation. The former results are reported in the above table as these were more favorable to the program.

[g] Data from MacKinnon et al. [19, p. 168, column 2].

[h] Dwyer et al. [20] used a number of different models in their analysis and concluded that these showed "mixed evidence of an effect on marijuana use" [20, p. 781). This is interpreted as a statistically significant positive effect on marijuana use (i.e., "✓" appears in the table).

[i] Data for the 1-year follow-up are from Pentz et al. [21, table 1]. The analysis of the 2-year follow-up data does not present exact cut-offs for *p* values (e.g., *p* <0.05). Rather it is stated that the values for cigarette use, marijuana use and alcohol use each fall between <0.05 and 0.001 (using 1-tailed tests) and a figure is presented that shows prevalence of use in the last month [21, figure 1]. One cannot determine from this if the difference reported for marijuana use would be statistically significant at *p* <0.05 (using a 2-tailed test). However, but since the effect at 1-year was said to be "maintained" [21, pp.220-221], a "?" appears in the table.

[j] Data from Johnson et al. [24, table 1].

modeling demonstrated such a meditational effect, which the authors interpreted as supporting "…the role of early adolescent drug use prevention programs in impacting later mental health problems" [9]. Riggs et al. [17] used the same methods to assess the effects of the MPP on marijuana use in early adulthood (assessment 12) and the subsequent effects of this on parent-child relationships and child impulsivity (assessment 14 or assessment 15). By this time the sample was reduced to just 541 subjects, of whom 257 had children in the age range for inclusion in the study. The analysis showed that the 131 subjects assigned to the MPP condition used significantly less marijuana in early adulthood compared to the 126 subjects in the control group and that there was a direct relationship from parental marijuana use to child impulsivity as well as an indirect relationship through parent-child interactions. Despite the fact that these results were produced from data obtained from just 16% of the subjects originally included in the MPP panel sample, Riggs et al. conclude that they "can be interpreted as the sustained impact of the MPP on adult marijuana use" and as "…suggesting that the MPP intervention contributed to lower levels of impulsivity in the children of original participants" [17].

Each of these later publications presents positive findings that the authors interpret as lending support to the idea that the MPP is an efficacious drug prevention program. However, each begs the question as to why these particular analyses were presented. A number of other "hard" drugs were assessed in the MPP evaluation, including cocaine, barbiturates and heroin [10]; so why do Riggs et al. [29] only focus on amphetamines? Use of alcohol could mediate the effects of an intervention program on receipt of mental health services as an adult or on parent-child interactions and child impulsivity; so why do Riggs and Pentz [9] and Riggs et al. [17] only focus on the potential meditational effects of marijuana? Why examine the effects of baseline marijuana use in one of these studies [17] but marijuana use during high school in the other [9]? Why use measurement points 13 and 14 to assess the outcome in one study [9] but 12 and 14 or 15 in the other [17]? Without clear answers to such questions one cannot be sure that the results presented in these recent publications from the Kansas City evaluation of the MPP are not analysis-driven rather than the result of by *a priori* hypothesis testing.

A Final Example from the Kansas City Evaluation

ISSN: 2330-2178

The analysis presented so far suggests that there has been a great deal of analytic flexibility in the reporting-out of results from data collected in the Kansas City MPP evaluation. A further illustration of this practice can be found in another paper published in the mid-2000s that used data from the evaluation to examine the effects of parent participation in the parent education components of the MPP on their perceptions of their influence over their children's drug use at the 2-year follow-up [30]. The data used in the analysis reported in this study came from a group of 351 parents of the 1,607 students in the panel sample. This group comprised just 28% of the 1,267 parents eligible to take part in the study at baseline. The analysis presented showed that the 233 parents who reported being in the MPP parenting program had significantly greater perceived influence over their children's drug use than the 118 parents who reported no participation, and this result was said to have "provided support for the effectiveness of the parent component of the MPP in changing parent perceptions of their children's substance use" [30].

As with the other more recent analyses that have emerged from the MPP Kansas City evaluation those presented by Riggs et al. [30] pertaining to involvement in the MPP parenting program raise numerous questions. For example, why are data reported for just the 2-year follow-up, and not the 1-year or 3-year follow-ups? Why is perceived parental control over adolescent drug-use the outcome variable and not the actual drug use of the participants' children who were in the MPP study? But by far the greatest indicator of analytic flexibility in this particular analysis is the manner in which the investigators construct the "MPP group" and the "control" group. As noted above, the panel sample of the MPP evaluation involved 1,607 children (and their parents) from eight schools that were randomly allocated to either receive the program (n=904) or not (n=703). That is, the investigators used standard research practices to create an intervention group and a control group. However, the analysis presented by Riggs et al. [30] completely abandoned this design: any parent who reported participation (regardless of whether he/she was from the intervention group or the control group) was included in the MPP group and was compared to any parent who did not report participation (regardless of whether he/she was from the intervention group or the control group). This would seem to be a textbook example of what Ioannidis [1] calls flexibility in study design.

The Indianapolis Experimental Evaluation of the Midwestern Prevention Project

The MPP Indianapolis replication study began in 1987 and involved 3,412 students in 6th and 7th grade at baseline (1,904 in the MPP group and 1,508 in the control group) [31]. No full account of the results of the Indianapolis experiment has been published. The data have been used to examine predictors of attrition over the course of the follow-up [32], parental substance use as a modifier of adolescent use [33], the characteristics of different types of community coalitions for alcohol and drug prevention [34], and the effects of parental participation in the parenting component of the program on children's drug use [35], but the effects of student participation in the main school-based component of the MPP on subsequent drug use are largely unknown.

The only published account to examine subsequent drug use of participants in the MPP Indianapolis replication focused on the "secondary" prevention effects of the program on those individuals who were already using cigarettes (n=400; 12% of the total sample), alcohol (n=613; 18% of the total sample) or marijuana (n=60; 1.8% of the total sample) in the previous month at baseline [31]. Subjects were followed-up at 6, 18, 30 and 42 months and statistically significant differences between the MPP group and the control group were observed at two of the four follow-ups for both previous-month cigarette use (the 6- and 30-month follow-up) and previous-month alcohol use (the 6- and 18-month follow-ups). This led the authors to conclude "…that social influence-based primary prevention programs can have an impact on not only students who are nonusers at baseline but also those who have began to use drugs" [31]. This conclusion largely overlooks the fact that the effects on alcohol use were short-lived and that there were no effects on marijuana use. Moreover, there is nothing in the MPP Indianapolis replication study to indicate that it has any effect on the drug use of baseline nonusers (i.e., the vast majority of those who took part in the study).

Conclusions

The above review of the results presented in publications from the two evaluations of the MPP shows that there is little evidence to indicate that the program had anything other than short-term effects cigarette use. With regard to the specific issue of flexibility in study design, definitions, outcomes and analytic models the focus of the Kansas City MPP evaluation began to narrow after the 2-year follow-up both in terms of the subjects included in the analyses and the outcomes for which results were reported. This tendency became even more noticeable in the publications that used data from the later follow-up points of the study when subjects from the panel sample were in their twenties and thirties. Given the scope of information collected and the number of assessment points, the Kansas City MPP evaluation must contain a huge amount of data about participants'

Table 4: Variables and Assessment Points Included in Recent Publications from the Kansas City MPP Evaluation Study.

		Assessment Point															
		Middle School			**High School**				**Emerging Adulthood**					**Early Adulthood**			
	Sample	B/L	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Riggs et al. [29]	1002	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
		Growth in amphetamine use through high school, emerging adulthood and early adulthood controlling for an index of baseline cigarette, alcohol and marijuana use.															
Riggs & Pentz [9]	961	✓			✓	✓	✓	✓							✓	✓	
		Marijuana use of subjects assessed at baseline and throughout high school. Receipt of mental health services assessed at age 27 and age 28-30 (measurement points 13 and 14).															
Riggs et al. [17]	257	✓												✓		✓	
		Marijuana use of subjects (parents) assessed at baseline and age 26 (measurement point 12). Child-parent interactions and child impulsivity assessed at age 28-30 or 30-34 (measurement point 14 or 15).															

ISSN: 2330-2178

drug use over more than a 20-year period, and yet just a few results have been reported. While such exploratory analyses are permissible, the results they produce should not be used (as they have been by the MPP evaluators) to make claims about the efficacy of the MPP as a prevention program. The piecemeal reporting of findings is even more noticeable in the Indianapolis replication study which has never published a paper that describes the effects of the program on the primary outcomes of cigarette, alcohol and marijuana use among the full sample. Commenting on this situation, Brian Flay (one of the evaluators on the MPP Kansas City study) observed that "…the lack of reported results from the Indianapolis replication, where the research design was stronger, leaves many questions unanswered and reflects negatively on the MPP study" [36].

It should be noted that such selective reporting of study findings within drug prevention research is not isolated to the MPP evaluation studies. Indeed, as Holder [2] warned, it is in danger of becoming normative within the field. It has been observed in the evaluations of the Life Skills Training program [3,5,37,38] the Seattle Social Development Project [5], Project ALERT [2,4,6] and the Strengthening Families Program [7,39]. Interestingly, it tends not to be found in independent replications of developer-led evaluations of drug prevention programs [e.g., 40,41]. This suggests that one condition that must be met before any program is designated "evidence-based" is that there be a truly independent replication study conducted by individuals with no ties to the program developers.

Ioannidis [1] observes that flexibility and selectivity in reporting of study findings is likely to be reduced by adherence to established standards for conducting and reporting controlled studies (e.g., those contained in the CONSORT statement [42]). Unfortunately, many of the journals that publish evaluation studies of drug prevention programs do not require registration with CONSORT. In addition, simply including a CONSORT style diagram of the follow-up of subjects over the course of a study (as was done in Riggs et al. [17,29]) does little to ensure the integrity of a study design, as it does nothing to prevent the selective reporting of study findings. Ultimately journal editors and reviewers are the key gatekeepers of research integrity within a field of inquiry and they must ensure that the claims made on behalf of a particular program are reasonable given the research design and analytic methods used to evaluate it and the outcomes reported. The MPP evaluation is one of the best know in the field of drug prevention studies since it was one of the earliest of the social influence programs to be studied, the findings from the early part of the study appeared in high profile journals, and the program has appeared on a number of evidence-based lists. Given this, it seems strange that those who reviewed the most recent publications from the Kansas City study would have been unaware that the results reported were selective and, in two cases, involved just a fraction of the subjects originally recruited into the study. In addition, it seems unlikely that the reviewers and editors would not have realized that the results presented in these papers were essentially the product of exploratory analyses that could not form the basis for making claims about the efficacy of the MPP. Again, suggests that confirmation bias is part of the general culture of the field of drug prevention research. Unless this changes, it is a field that will have no sound scientific base [2].

References

1. Ioannidis JPA (2005) why most published research findings are false. PLoS Med 2: e124 (0696-0701).
2. Gerstein DR, Green LW (1993) Preventing drug abuse: what do we know? Washington, DC: National Academy Press.
3. Brown JH (2001) Youth, drugs and resilience education. J Drug Educ 31: 83-122.
4. Gandhi AG, Murphy-Graham E, Petrosino A, Chrismer SS, Weiss CH (2007) The devil is in the details: examining the evidence for "proven" school-based drug abuse prevention program. Eval Rev 31: 43-74.
5. Gorman DM (2005) Drug and violence prevention: Rediscovering the critical rational dimension of evaluation research. J Exp Criminol 1: 39-62.
6. Gorman DM, Conde E (2010) The making of evidence-based practice: The case of Project ALERT. Children Youth Serv Rev 32: 214-222.
7. Gorman DM, Conde E, Huber Jr JC (2007) The creation of evidence in "evidenced-based" drug prevention: A critique of the Strengthening Families Program Plus Life Skills Training evaluation. Drug Alcohol Rev 26: 585-593.
8. Holder HD (2010) Prevention programs in the 21st century: What we do not discuss in public. Addiction 105: 578-581.
9. Riggs NR, Pentz MA (2009) Long-term effects of adolescent marijuana use prevention on adult mental health services utilization: the Midwestern Prevention Project. Subst Use Misuse 44: 616-631.
10. Pentz MA, Mihalic SF, Grotpeter JK (1997) Blueprints for violence prevention: the Midwestern Prevention Project. Boulder, Colorado: Institute of Behavioral Science, University of Colorado.
11. Mihalic S, Irwin K, Elliot D, Fagan A, Hansen D (2001) Blueprints for violence prevention. Washington, DC: US Department of Justice.
12. National Registry of Evidence-based Programs and Practices (2013) Legacy programs: Project STAR: Students Taught Awareness and Resistance.
13. National Institute on Drug Abuse (2003) Preventing drug use among children and adolescents: A research-based guide for parents, educators, and community leaders (2nd Edition), Bethesda, MD: US Department of Health and Human Services.
14. Promising Programs Network (2011) Programs that work: Midwestern Prevention Project/Project STAR.
15. US Department of Health and Human Services (2001) Youth violence: A report of the Surgeon General. Rockville, MD: US Department of Health and Human Services.
16. Pentz MA, Dwyer JH, MacKinnon DP, Flay BR, Hansen WB, et al. (1989) Multicommunity trial for primary prevention of adolescent drug use: Effects on drug use prevalence. JAMA 261: 3259-3266.
17. Riggs NR, Chou C-P, Pentz MA (2009) Protecting against intergenerational problem behavior: mediational effects of prevented marijuana use on second-generation parent-child relationships and child impulsivity. Drug Alcohol Depend 100: 153-160.
18. Pentz MA, Trebow EA, Hansen WB, MacKinnon DM, Dwyer JH, et al. (1990) Effects of program implementation on adolescent drug use behavior: The Midwestern Prevention Project. Eval Rev 14: 264-289.
19. MacKinnon DP, Johnson CA, Pentz MA, Dwyer JH, Hansen WB, et al. (1991) Mediating mechanisms in a school-based drug prevention program: First-year effects of the Midwestern Prevention Project. Health Psychol 10: 164-172.
20. Dwyer JH, MacKinnon DP, Pentz MA, Flay BR, Hansen WB et al. (1989) Estimating intervention effects in longitudinal studies. Am J Epidemiol 130: 781-795.
21. Pentz, M.A., Johnson CA, Dwyer JH MacKinnon DM, Hansen WB, et al. (1989) A comprehensive community approach to adolescent drug abuse prevention: Effects on cardiovascular disease risk behaviors Ann Med 21: 219-222.
22. Pentz, M.A., MacKinnon, D.P., Dwyer, J.H., Wang, E.W.I., Hansen, et al. (1989) Longitudinal effects of the Midwestern Prevention Project on regular and experimental smoking in adolescents. Prev Med 18: 304-321.

ISSN: 2330-2178

23. Pentz MA, MacKinnon DP, Flay BR, Hansen WB, Johnson CA, et al. (1989) Primary prevention of chronic diseases in adolescence: effects of the Midwestern Prevention Project on tobacco use. Am J Epidemiol 130: 713-724.

24. Johnson CA, Pentz MA, Weber MD, Dwyer JH, Baer N, et al. (1990) Relative effectiveness of comprehensive community programming for drug abuse prevention with high-risk and low-risk adolescents. J Consult Clin Psychol 58: 447-456.

25. Ringwalt C, Paschall MJ, Gorman DM, Derzon J, Kinlaw A (2011) The use of one- versus two-tailed tests to evaluate prevention programs. Eval Health Prof 34: 135-150.

26. Valente TW, Ritt-Olson A, Stacey A, Unger JB, Okamoto J, et al. (2007) Peer acceleration: effects of a social network tailored substance abuse prevention program among high-risk adolescents. Addiction 102: 1804-1815.

27. Ellickson PL, Bell RM (1990) Drug prevention in junior high: a multi-site longitudinal test. Science 247: 1299-1305.

28. Hallfors D, Cho H, Sanchez V, Khatapoush S, Kim HM, et al. (2006) Efficacy vs. effectiveness trial results of an indicated "model" substance abuse program: Implications for public health. Am J Public Health 96: 2254–2259.

29. Riggs NR, Chou CP, Pentz MA (2009) Preventing growth in amphetamine use: long-term effects of the Midwestern Prevention Project (MPP) from early adolescence to early adulthood. Addiction 104: 1691-1699.

30. Riggs NR, Elfenbaum P, Pentz MA (2006) Parent program component analysis in a drug abuse prevention trial. J Adolesc Health 39: 66-72.

31. Chou CP, Montgomery SB, Pentz MA, Rohrbach LA, Johnson CA, et al. (1998) Effects of a community-based prevention program on decreasing drug use in high-risk adolescents. Am J Public Health 88: 944-948.

32. Fan AXZ, Pentz MA, Dwyer J, Bernstein K, Li C (2002) Attrition in a longitudinal drug abuse prevention study. Int J Mental Hlth Prom 4: 15-23.

33. Li C, Pentz MA, Chou CP (2002) Parental substance use as a modifier of adolescent substance use risk. Addiction 97: 1537-1550.

34. Mansergh G., Rohrbach LA, Montgomery SB, Pentz MA, Johnson CA (1996) Process evaluation of community coalitions for alcohol and other drug abuse prevention: a case study comparison of researcher- and community-initiated models. J Comm Psychol 24: 118-135.

35. Rohrbach LA, Hodgson CS, Broder BI, Montgomery SB, Flay BR, et al. (1994) Parental participation in drug abuse prevention: results from the Modwestern Prevention Project. J Res Adolescence 4: 295-317.

36. Flay BR (2000) Approaches to substance use prevention utilizing school curriculum plus social environment change. Addict Behavs 25: 861-885.

37. Gorman DM (2005) Does measurement dependence explain the effects of the Life Skills Training Program on smoking outcomes? Prev Med 40: 479-487.

38. Gorman DM (2011) Does the Life Skills Training program prevent use of marijuana? Addiction Res Theory 19: 470-481.

39. Gorman DM, Conde E (2009) Further comments on the path to drawing reasonable conclusions about prevention. Addiction 104: 152-154.

40. Hallfors D, Cho H, Sanchez V, Khatapoush, S, Kim HM, et al. (2006) Efficacy vs effectiveness trial results of an indicated "model" substance abuse program: implications for public health. Am J Public Health 96: 2254-2259.

41. Ringwalt CL, Clark HK, Hanley S, Shamblen SR, Flewelling RL (2010) The effects of Project ALERT one year past curriculum completion. Prev Sci 11: 172-184.

42. Moher D, Schulz KF, Altman DG (2001) The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 357: 1191–1194.