

Prevencción del consumo de sustancias: la absoluta necesidad de una evaluación aleatorizada u observacional

Prevention of substance abuse: Randomised or observational evaluation is absolutely needed

FABRIZIO FAGGIANO

Università del Piemonte Orientale "A.Avogadro"
Dipartimento di Medicina Clinica e Sperimentale

Enviar correspondencia a:
Fabrizio Faggiano
Università del Piemonte Orientale "A.Avogadro"
Dipartimento di Medicina Clinica e Sperimentale
Via Solaroli, 17 – 28100 Novara
Tel. +39 0321/660.661-682 · Fax. +39 0321/660.682
E-mail: fabrizio.faggiano@med.unipmn.it

RESUMEN

Los programas e intervenciones preventivas en el campo del consumo de drogas buscan reducir el consumo, pero a veces pueden ser contraproducentes. Por ello se necesitan evaluaciones rigurosas. En Europa, a pesar de la amplia difusión de dichas intervenciones, casi nunca se evalúan con métodos rigurosos.

Por esa razón es importante promover la evaluación de los programas que se consideren relevantes. Así, otros profesionales podrán elegir intervenciones basándose en la evidencia de su eficacia.

Pero, ¿qué tipo de evaluación de programas de prevención se necesitan? Existe un amplio consenso sobre el papel de la evaluación aleatorizada. La aleatorización es el único medio eficaz de controlar los numerosos factores que pueden sesgar los resultados. El modelo de evaluación típico asociado a la mayoría de los programas de intervención, la evaluación pre-post, no efectúa valoraciones imparciales.

Sin embargo, hay muchas intervenciones que no se pueden aleatorizar porque no existen a nivel individual (por ejemplo, las políticas sobre el precio del tabaco). Estos programas exigen el método observacional para su evaluación, como en una cohorte o en un modelo de series temporales interrumpidas. Aunque estos tipos de estudio tienen una metodología compleja, la evaluación es de gran interés para el desarrollo de estrategias de prevención basadas en la evidencia.

Palabras clave: *prevención drogas, evaluación, aleatorización, métodos observacionales*

ABSTRACT

Prevention programs and interventions in the field of drug abuse are designed to reduce the incidence of onset of substance use, but they can also produce harms. This is the reason why rigorous evaluations are absolutely needed. In Europe, in spite of the large spread of such interventions, they are rarely evaluated with rigorous study designs.

For these reasons it is important to promote the evaluation of every program that can be considered of relevance. This would allow the other practitioners to choose the interventions on the base of evidence of effectiveness.

But which kind of evaluation prevention programs needed? There is a substantial agreement on the role of the randomized evaluation. Randomization is the only effective way to control for numerous factors that can bias the results. The typical evaluation design associated to most part of intervention programs, the pre-post assessment, cannot produce unbiased estimations.

There are however many interventions that cannot be randomised, because they cannot be provided at the individual level (e.g like the price policies for cigarettes). These programs require an observational approach for evaluation, like a cohort or an Interrupted Time Series design. Even if these study designs have a complex methodology, the evaluation is of high relevance in order to allow the development of prevention strategies based on evidence.

Key words: Drug prevention, evaluation, randomization, observational methodology.

Un reciente informe sobre medidas contra las drogas ilegales en Europa demuestra que las intervenciones preventivas están muy difundidas, pero la mayoría no se apoyan en una evidencia de eficacia científica¹. Entonces, ¿cuál es el fundamento en que se basan esas intervenciones? Seguramente, la teoría.

La literatura científica presenta numerosos casos que indican que la mera teoría no basta para tomar buenas decisiones. La literatura médica está llena de ejemplos de intervenciones basadas en la teoría que, al someterse a evaluaciones rigurosas, mostraron más perjuicios que beneficios^{2,3}. Pero hay un par de contundentes ejemplos en el campo de la prevención del consumo de sustancias que inducirán a los profesionales a meditar a fondo su trabajo.

El primero es la *American National Youth Anti-drug Media Campaign*: esta campaña fue elaborada en 1997 por la *US National Drug Control Policy* (ONDCP) con el objetivo primordial de «educar y adiestrar a los jóvenes estadounidenses para que rechacen las drogas ilegales, el alcohol y el tabaco», aunque el alcohol y el tabaco dejaron de ser prioritarios posteriormente⁴. La campaña se diseñó según las teorías más fiables y el Congreso de Estados Unidos la financió generosamente (¡mil quinientos millones de dólares!). Consistió, sobre todo, en anuncios televisivos de sensibilización contra las drogas, emitidos por las cadenas más importantes desde 1998 a 2004. Como gran parte de las políticas importantes en Estados Unidos, la campaña se evaluó para valorar sus efectos sobre la incidencia y prevalencia del consumo de marihuana. La evaluación no mostró evidencias de efectos positivos en el consumo de drogas entre los adolescentes, pero sí presentó indicaciones de impacto negativo. En concreto, el consumo mensual de marihuana creció significativamente en el grupo entre 14-18 años (riesgo relativo ajustado [RR]=1,21; intervalo de confianza al 95% [IC] 1,19-1,65), y los que no respondieron «rotundamente no» al consumo de marihuana descendieron de forma notoria (RR ajustado = 0,89; IC al 95% 0,79-1,00).

El segundo ejemplo es el *Adolescent Substance Abuse Prevention Study* (ASAPS)⁵. Se llama así un amplio estudio controlado aleatorio (RCT) en el que participaron 20.000 estudiantes para evaluar un programa de prevención basado en la escuela y denominado *Take Care of Your Life –Cuida tu vida–* (TCYL). El TCYL, que utilizó los principios de programas exitosos de prevención de consumo de sustancias, consistía en 10 clases en el séptimo curso (1º de ESO) y 7 clases de refuerzo en noveno curso (3º de ESO). El estudio de evaluación contempló un amplio RCT de grupos que abarcó varias ciudades estadounidenses. Los resultados fueron abrumadores: el consumo mensual de alcohol aumentó significativamente un 9% (RR=1,09; IC al 95% 1,01-1,18), la frecuencia mensual de borracheras aumentó el 10% (RR=1,10; IC al 95% 0,98-1,22), y el consumo mensual de cigarrillos aumentó el 21% (RR=1,21; IC al 95% 1,05-1,37). El único resultado positivo fue una pequeñísima reducción en el consumo mensual de marihuana.

Ambos casos son evaluaciones rigurosas de intervenciones preventivas basadas en las mejores teorías, elaboradas y desarrolladas por prestigiosos investigadores, y bien dotadas

A recent report on illicit drugs responses in Europe confirms that the prevention interventions are largely disseminated, but that most part of them are not supported by scientific evidence of effectiveness¹. What's, therefore, the decision-base for these interventions? Probably, is theory.

The scientific literature present many cases suggesting that theories alone are not reliable in order to take good decisions. The entire medical literature is plenty of examples of theory-based interventions that, once submitted to rigorous evaluations, showed more harms than benefits^{2,3}. But there are a couple of striking examples from the field of substance abuse prevention that should encourage practitioners to reflect on their practice.

The first is the *American National Youth Anti-drug Media Campaign*: this campaign as been planned in 1997 by the *US National Drug Control Policy* (ONDCP) with the main objective "to educate and enable America's youth to reject illegal drugs as well as alcohol and tobacco", even if alcohol and tobacco were later omitted from the main focus of the campaign⁴. The campaign was designed applying the most reliable theories and adequately funded by the Congress of the United States (\$1.5 billion dollars !). It was based on televised antidrug public service announcements (PSAs), broadcasted in all the most important TV networks from 1998 to 2004. As for most part of relevant US policies, the campaign was evaluated in order to assess its effects on the incidence and prevalence of marijuana use. Evaluation provided no evidence of positive effect in relation to teen drug use, and showed some indications of a negative impact. In particular, the past month use of marijuana significantly increased among 14-18 years (adjusted RR=1.21; 95%CI 1.19-1.65), and those who did not answer "definitely not" on intention to use marijuana significantly decreased (adjusted RR=0.89; 95%CI 0.79-1.00).

The second example is the *Adolescent Substance Abuse Prevention Study* (ASAPS)⁵. This is the name of a large Randomised Controlled Study (RCT), involving about 20,000 students, to evaluate a school-based prevention program called *Take Care of Your Life* (TCYL). TCYL used the principles of successful substance abuse prevention programs, and consisted of 10 lessons in the 7th grade of school and of 7 booster lessons in 9th grade. The evaluation study was a large Cluster RCT, involving several US cities. The results were impressive: the 30-day use of alcohol significantly increased by 9% (RR=1.09; 95%CI 1.01-1.18), the 30-day frequency of drunkenness increased by 10% (RR=1.10; 95%CI 0.98-1.22) and the 30-day use of cigarettes increased by 21% (RR=1.21; 95%CI 1.05-1.37). The only positive result was a not significant reduction in 30-day marijuana use.

Both cases are rigorous evaluations of preventative interventions based on the best available theories, designed and carried out by really authoritative researchers, well

económica y administrativamente. ¡Pero provocaron perjuicios en vez de beneficios!

De estos casos se desprenden tres lecciones:

- i. La teoría, las buenas intenciones, los profesionales de renombre y una generosa financiación no bastan para predecir el efecto de una intervención preventiva. En disciplinas con altos niveles de complejidad, como la biología, la psicología o las ciencias sociales, las teorías son vacilantes y poco fidedignas hasta que las avala una evidencia científica consistente⁶.
- ii. La extensión de los programas preventivos no es aceptable desde un punto de vista ético si no cuenta con una evaluación rigurosa. Cuando un profesional, con las mejores intenciones, decide desarrollar un programa de prevención, tiene que saber a ciencia cierta si la intervención resultará beneficiosa⁷. Y la única forma de saberlo es basar su decisión en evidencias de eficacia.
- iii. La evaluación de los programas y la publicación de sus resultados en revistas científicas es deber de todos los promotores de intervenciones preventivas. De ese modo, los demás profesionales podrán elegir intervenciones basándose en evidencias de eficacia. Éste es el papel, por ejemplo, del *Best Practice Portal* (Portal de Buenas Prácticas) del EMCDDA (<http://www.emcdda.europa.eu/themes/best-practice/evidence/selective-prevention>). Los profesionales e investigadores deben mostrarse agradecidos sobre todo con los autores que publican resultados negativos, puesto que su honradez científica muchas veces es más útil para la salud pública que la publicación de resultados positivos.

Se discute la evaluación de los programas de prevención⁸, pero el papel de la evaluación aleatorizada goza de gran respaldo. La aleatorización es la única forma eficaz de controlar los numerosos factores susceptibles de sesgar los resultados, sesgos que por lo general favorecen a los partidarios de la intervención. El típico modelo de evaluación asociado a la mayoría de los programas de intervención, la valoración pre-post, no ofrece información imparcial sobre la eficacia de los programas. Sin embargo, combinado con los resultados de un RCT realizado en una población distinta es muy útil para el proceso de evaluación, la fidelidad al programa original y la valoración de las variables intermedias con el fin de predecir su eficacia en el contexto específico.

Y así, tanto en la prevención como en el terreno de la evaluación del tratamiento, el RCT es la clave ideal para cualquier intervención que se proporcione a nivel individual y que, por ello, se pueda aleatorizar. En intervenciones de grupo, como los programas escolares, en los que al menos el grupo se puede aleatorizar, la referencia básica es el RCT de grupos⁹.

Pero hay muchas intervenciones que no se ofrecen a nivel individual ni de grupo, como la política de precios del tabaco y de las bebidas alcohólicas, las campañas de los medios de comunicación, las prohibiciones o las restricciones de consumo. Se trata de intervenciones con un efecto esperado

funded and supported. But they caused harm, instead of benefit!

There are three lessons to learn from these cases:

- iv. theory, as well as good intentions, authoritative professionals, adequate funds, cannot predict alone the effect of a preventative intervention. In fields with high levels of complexity, like biology, psychology or social sciences, every theory is unsteady and unreliable until supported by strong scientific evidence⁶;
- v. the dissemination of prevention programs, if not endowed with a rigorous evaluation, is not acceptable from an ethical point of view. When a practitioner, with the better intentions, decide to carry out a prevention program, he has to be sure that the intervention will cause benefits⁷. The only way to be sure is to base his decision on evidence of effectiveness;
- vi. the evaluation of programs, and the publication of results in scientific reviews, is a duty of all the developers of preventative interventions. This would allow the other practitioners to choose the interventions on the base of evidence of effectiveness. This is the role, for example, of the Best Practice Portal of EMCDDA (<http://www.emcdda.europa.eu/themes/best-practice/evidence/selective-prevention>). All professionals and researchers have to be especially grateful to authors that publish negative results, because their scientific honesty can be more useful for public health than the publication of positive results.

There is some debate about evaluation of prevention programs⁸, but there is a large agreement on the role of the randomized evaluation. Randomization is the only effective way to control for numerous factors that can bias the results, and these biases usually favors the intervention arm. The typical evaluation design associated to most part of intervention programs, the pre-post assessment, cannot give any unbiased information about effectiveness of program. On the other side, it can be really useful, together with the results of a RCT eventually carried out in a different population, for evaluating process, fidelity on the original program, and also for assessing the role of intermediate variables in order to predict its effectiveness in the specific context.

Thus, in prevention as well as in the treatment evaluation field, RCT is the gold standard for evaluation of any intervention that can be provided at the individual level, and therefore randomised. For interventions provided at the level of group, like school-based programs, in which at least the group can be randomised, the Cluster RCT is the reference⁹.

But there are many interventions that cannot be provided at the individual level neither at the group level, like the price policies for cigarettes or alcohol beverages, mass media campaigns, bans or restrictions in use. These

de antemano en el problema de salud elegido, que se aplican ampliamente sin evaluación o no se aplican en absoluto por falta de evaluación. Estas intervenciones, en las que no se puede poner en práctica el paradigma del RCT, exigen un enfoque de evaluación no aleatorizado. Por ejemplo, las campañas de los medios de comunicación se pueden evaluar según un diseño de cohorte, como la evaluación de la *American National Youth Anti-drug Media Campaign*⁴, mientras que las políticas de precios o las prohibiciones se evalúan según un enfoque de cohorte¹⁰, o según series temporales interrumpidas¹¹. Estos modelos de estudio son especialmente complejos debido al riesgo de parcialidad que entraña el diseño observacional¹². Por ejemplo, las diferencias de problemas asociados al alcohol entre diversas poblaciones se explican a veces por las diferencias de precio de las bebidas alcohólicas¹², pero hay que tener en cuenta otros muchos factores, como las diferencias culturales contra el alcohol en el marco de las creencias religiosas, que llevan a una valoración errónea de la importancia del precio. Cuando sea posible, a saber, cuando la intervención pueda ser aleatorizada, debe recurrirse al RCT para evitar la parcialidad. Cuando no queda más remedio que optar por un estudio no aleatorizado, hay que prestar especial atención al diseño e interpretación del mismo, teniendo en cuenta la fragilidad intrínseca de la observación.

El desarrollo de estrategias de prevención basadas en la evidencia es esencial para mejorar su eficacia comunitaria, evitando la elección de intervenciones ineficaces y a veces incluso perjudiciales. Para ello los investigadores deben ofrecer pruebas rigurosas de prácticas de intervención eficaces no susceptible de aleatorización, y los responsables de la administración han de exigir a los investigadores las evidencias necesarias para tomar decisiones efectivas.

Por su parte, investigadores y profesionales han de jugar el difícil papel de desaconsejar el empleo de prácticas no evaluadas y de alentar el de prácticas basadas en la evidencia.

are interventions with a large expected effect on the target health problem, which are either extensively applied without evaluation or not applied at all because of lack of evaluation. These interventions, where the paradigm of RCT cannot be applied, require a non-randomised approach for evaluation: mass media campaigns, for example, can be evaluated by a cohort design, like the evaluation of the American 'National Youth Anti-drug Media Campaign'⁴, whereas price policies or bans can be evaluated by cohort approach¹⁰, or by Interrupted Time Series¹¹. These study designs are particularly complex because of the risk of bias that the observational design entails¹². For example, differences in alcohol-related problems among different populations can be explained by differences in price of alcoholic beverages¹², but many other factor could be implied, as for example differences in culture against alcohol, in level of religiousness, conducting to wrong estimation of the role of price. Whenever possible, that is, when the intervention can be randomised, RCT have to be chosen in order to avoid bias. When the choice of a non-randomised study is forced, a particular attention have to be put in the design and interpretation of the study, being aware of the intrinsic weakness of the observation.

The development of prevention strategies based on evidence is essential in order to improve their community effectiveness, avoiding the choice of ineffective and, sometimes, harmful interventions. This requires that researchers have to learn how to better provide rigorous evidence of effective practices for interventions that cannot be randomised, and policy-makers have to learn to pretend by researchers the evidence needed for effective decision-making.

On the other hand researchers and practitioners have to play the difficult role of discouraging the use of practices not evaluated, and of encourage the use of evidence based practices.

REFERENCIAS /REFERENCES

1. EMCDDA. Annual report 2009. EMCDDA. Lisboa, 2009.
2. Bjelakovic G, Nikolova D, Gluud LL, Simonetti RG, Gluud C. Mortality in Randomized Trials of Antioxidant Supplements for Primary and Secondary Prevention - Systematic Review and Meta-analysis. *JAMA* 2007; 297: 842-857.
3. Women's Health Initiative writing group. Risks and benefits of estrogens plus progestin in healthy postmenopausal women. *Jama* 2002; 288: 321-333.
4. Hornik R. Personal Influence and the Effects of the National Youth Anti-Drug Media campaign. *The ANNALS* 2006; 608: 282-300.
5. Sloboda Z, Stephens RC, Stephens PC, Grey SF, Teasdale B, Hawthorne RD, Williams J, Marquette JF. The Adolescent Substance Abuse Prevention Study: A randomized field trial of a universal substance abuse prevention program. *Drug Alcohol Depend* 2009; 102: 1-10.
6. Popper K. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge Classics 1994, Londres.
7. Chalmers I. Trying to do more Good than Harm in Policy and Practice: The Role of Rigorous, Transparent, Up-to-Date Evaluations. *The ANNALS* 2003; 589: 22
8. Victora CG, Habicht J-P, Bryce J. Evidence-Based Public Health: Moving Beyond Randomized Trials. *Am J Public Health* 2004; 94: 400-5.
9. Faggiano F, Vigna-Taglianti F, Burkhart G, Bohrn K, Cuomo L, Gregori D, Panella M, Scatigna M, Siliquini R, Varona L, van der Kreeft P, Vassara M, Wiborg G, Galanti MR; the EU-Dap Study Group. The effectiveness of a school-based substance abuse prevention program: 18-Month follow-up of the EU-Dap cluster randomized controlled trial. *Drug Alcohol Depend* 2010. [publicación en Internet previa a edición]
10. Barone-Adesi F, Vizzini L, Merletti F, Richiardi L. Short-term effects of Italian smoking regulation on rates of hospital admission for acute myocardial infarction. *Eur Heart J* 2006; 27:2468-72.
11. Chaloupka FJ, Grossman M, Saffer H. The effects of price on alcohol consumption and alcohol-related problems. *Alcohol Res Health* 2002; 26-34.
12. Cook DT, Campbell DT. *Quasi-experimentation: design and analysis issues for field settings*. Chicago, Rand McNally, 1979.