International Guidelines for the Estimation of the Avoidable Costs of Substance Abuse

By:

David Collins
Helen Lapsley
Serge Brochu
Brian Easton
Augusto Pérez-Gómez
Jürgen Rehm
Eric Single

1st edition - November 2006

Preface

These guidelines were commissioned by Health Canada as part of an international initiative to develop sound methodologies and approaches for estimating the socio-economic avoidable costs of substance abuse. Health Canada would like to extend its appreciation and thanks to the participants of this initiative who are each recognized in their own country, as an expert in the field.

The current document is meant to provide guidance for developing pilot studies on estimating avoidable costs. Because it is expected that these guidelines will continually evolve with future applications and studies, an *International Steering Committee on Estimating Avoidable Costs of Substance Abuse* was established. Governments and organizations that plan to undertake pilot studies on estimating avoidable costs of substance abuse are invited to become members of the International Steering Committee.

It is hoped that these guidelines will be helpful to both developing and developed countries. When undertaking studies, guideline users are strongly advised to focus on a single substance, e.g. alcohol, tobacco, or illicit drugs. As well, before *avoidable* costs can be estimated, good basic data on *aggregate* costs of the substance being studied must already exist.

For questions or information regarding the guidelines and the *International Steering Committee on Estimating Avoidable Costs of Substance Abuse*, please contact Health Canada at the following address.

Office of Research and Surveillance
International Steering Committee on
Estimating Avoidable Costs of Substance Abuse
Drug Strategy and Controlled Substances Programme
Healthy Environments and Consumer Safety Branch
Health Canada
123 Slater Street, A.L. 3509C
Ottawa ON K1A 1B9
E-mail: ors_brs@hc-sc.gc.ca
Fax: (613) 948-7977

These guidelines are available in both English and in French. Copies are available for downloading from the Health Canada website (www.hc-sc.gc.ca).

Table of contents

List of tables

Acknowledgments

These guidelines are the result of collaborative work between members of the International Working Group, appointed to oversee the project, and participants at the June 2005 Ottawa workshop organised and funded by Health Canada. David Collins (Australia) and Helen Lapsley (Australia) wrote the original discussion document which was presented and reviewed at the Ottawa workshop. They then undertook extensive revision and expansion of the document informed by other workshop papers, discussions at the workshop and subsequent written comments from workshop participants. The second draft was circulated to the Working Group for further comment, following the receipt of which the final version was produced.

In such a collaboration, assigning credit to individual authors is not simple. However, some authors have made particularly significant and identifiable contributions:

- Jürgen Rehm (Canada and Switzerland), Benjamin Taylor (Canada), Jayadeep Patra (Canada) and Gerhard Gmel (Canada and Switzerland) for the exposition "From the classical to the distributional method in epidemiology";
- Brian Easton (New Zealand) and Eric Single (Canada) for the discussion of the reliability and usefulness of avoidable cost estimates;
- Serge Brochu (Canada) for discussion of the derivation of the attributable fractions for crime;
- Eric Single (Canada) for the examination of the implications of unrecorded alcohol production and consumption for the estimation of the social costs of alcohol;
- Augusto Pérez-Gómez (Colombia) for discussion of issues relating to the estimation of the avoidable costs of substance abuse in Central and Southern America.

While all workshop participants made valuable contributions, it should be acknowledged that particularly important perspectives came from (in alphabetical order) Serge Brochu (Canada), Brian Easton (New Zealand), Rick Harwood (United States), Claude Jeanrenaud (Switzerland), Pierre Kopp (France), Jacques Le Cavalier (Canada), Augusto Pérez-Gómez (Colombia), Jürgen Rehm (Canada and Switzerland) and Tim Stockwell (Canada).

The workshop, and indeed the whole project, would not have been possible without the invaluable contributions of Louise Déry and her team from Health Canada.

1 Introduction

1.1 Background

Between 1994 and 2002 the Canadian Centre on Substance Abuse (CCSA) held a series of symposia and workshops in Canada and the United States, with the purpose of developing a set of guidelines for estimating the social costs of substance abuse (in particular, alcohol, tobacco and illicit drugs). The guidelines were intended to encourage and facilitate international research, in both developed and developing countries, into the costs of substance abuse.

The meetings brought together international experts involved in this type of research together with policymakers, bureaucrats and NGO representatives interested in the policy applications of such research. As a result of these meetings a team of experts, mainly economists and epidemiologists, from a range of countries (Australia, Canada, Colombia, France, New Zealand and the USA) produced two editions of *International guidelines for estimating the costs of substance abuse*. These were originally published by the CCSA. The second edition, with minor modifications, was subsequently published by the World Health Organization (see Single *et al*, 2003) and, as a result, has achieved wide circulation. It has been influential in encouraging the spread of international research projects to estimate the social costs of substance abuse.

As the *Guidelines* explained, estimates of the costs of substance abuse constitute just one component in a range of potential economic information on substance abuse. The following table, from Single *et al*, presents a summary of the hierarchy of potential economic information on the costs of substance abuse, together with examples of their possible uses.

Table 1 – Substance abuse cost estimates and their policy uses

Type of estimate	Interpretation of results	Example of policy use
Aggregate costs	Total external costs of substance abuse compared with the alternative situation of no substance abuse	Indication of the size of the substance abuse problem
Avoidable costs	Potential economic benefits from substance abuse harm minimization strategies	Determination of the appropriate level of resources to be devoted to harm minimization strategies
Costs incidence	The distribution of the external costs of substance abuse among various community groupings	Mobilisation of support from various groups (for example, the business community) for anti-substance abuse programs
Disaggregated costs	External costs of substance abuse disaggregated by categories	Economic evaluation (cost-benefit or cost-effectiveness analysis) of harm minimization programs
Budgetary impact	The impact of substance abuse on government expenditures and revenues	Assessment of the case for industry compensation payments to government as a result of abusive use of substances which the industry produces

The *Guidelines* provide considerable information on estimation of what, in Table 1 above, are referred to as *aggregate costs*. However, while aggregate cost estimates are extremely valuable for a number of purposes, as an indicator of the overall economic burden borne by the community as a result of substance abuse, they indicate neither the proportion of such aggregate costs which are potentially avoidable nor the nature of the programmes and policies best suited to achieve this cost avoidance. It was agreed by participants at workshops convened to develop the original guidelines that the next logical step in this area of research would be to proceed to the estimation of avoidable costs, which indicate the benefits potentially available to harm minimization programs.

In 2004, at the 46th Session of the Commission on Narcotic Drugs, the Government of Canada informed member states of its intention to launch an international initiative to develop guidelines for estimating avoidable costs of substance abuse. CCSA proposed the convening of a workshop to develop these guidelines. Health Canada agreed to financially support this initiative and pulled together a group of international experts to develop the guidelines. An international workshop was held in Ottawa in June 2005. The

workshop attendees represented a wide range of academic disciplines, organizational backgrounds and countries. Two Australian economists, David Collins and Helen Lapsley, were commissioned to produce for the workshop a paper surveying the issues and problems involved in estimating the avoidable costs of substance abuse. Other papers dealing with specific areas of this research exercise were also commissioned for presentation at the workshop. The survey paper and the other commissioned papers formed the basis for discussions at the workshop. The present avoidable cost guidelines represent a revised version of the original survey paper, very substantially rewritten and extended by its original authors in the light of the other commissioned papers and of discussions at the workshop. Where these new avoidable cost guidelines make extensive use of, or reference to, workshop papers or the comments of workshop participants, their authors are also formally acknowledged as joint authors of this report.

It is intended, as far as possible, to avoid overlap between this study and the previously published *Guidelines* for estimating the aggregate social costs of substance abuse. This report concentrates on estimating the *avoidable* costs of substance abuse. It does not attempt to discuss the wider issues involved in estimating the aggregate social costs of substance abuse except insofar as they are relevant to the estimation of avoidable costs. It is presumed that readers will already have some familiarity with the original aggregate cost *Guidelines*.

However, international data on substance abuse have improved and, in some respects, the original *Guidelines* have the potential for further development. This is particularly true in areas such as the epidemiological evidence about the effects of substance abuse and in the identification of drug-attributable crime. Accordingly, it is hoped that a third edition of the aggregate cost *Guidelines* will eventually be produced. In the meantime, this new information needs to be acknowledged in the avoidable cost guidelines, even though technically it is more specifically related to the estimation of aggregate costs than to avoidable costs. Thus, the explanation of these developing areas is presented here but, in order to maintain the logical flow of this report, the information is presented in appendices rather than in the main body of the report.

The aggregate cost *Guidelines* were produced by a group of authors all of whom had practical experience in the estimation of the social costs of substance abuse. On the other hand, since the estimation of substance abuse *avoidable* costs is such a new area of research, in which there is little published literature, few of the present authors have practical experience in this research area. With avoidable cost projects in various countries likely to be stimulated by the publication of avoidable cost guidelines, it is to be expected that future editions will be able to be substantially revised in the light of the practical experience gained in these projects.

1.2 The nature of avoidable costs

In estimating the aggregate costs of substance abuse, comparison must be made between the actual substance abuse situation and some alternative hypothetical situation, known as the counterfactual situation. This counterfactual, though often not specifically identified, is in most cost studies that of a situation of no past or present abuse of the substance(s) in question. Thus the counterfactual situation is implicitly one in which the community would bear no substance abuse costs. By comparing the actual situation with the counterfactual zero-abuse situation, the extra costs which that abuse imposes on the community can be calculated.

However, the hypothetical counterfactual situation is, as Single *et al* point out, "hypothetical and not realizable under any circumstances. Estimates of the *total* costs of drug abuse comprise both avoidable and unavoidable costs. Unavoidable costs comprise the costs which are currently borne relating to drug abuse in the past, together with the costs incurred by the proportion of the population whose level of drug consumption will continue to involve costs. Avoidable costs are those costs which are amenable to public policy initiatives and behavioural changes."

Thus, avoidable cost estimates provide an indication of the benefits potentially available to the community as a whole as a result of directing public resources to the prevention or reduction of substance abuse. They provide valuable economic information on the basis of which a more efficient allocation of productive resources could be achieved. It would be theoretically possible to have a situation in which, although aggregate substance abuse costs were high, avoidable costs were so low as not to justify public expenditures designed to reduce abuse. In practice this is an unlikely scenario but it is one that might well be put forward, as an interpretation of aggregate cost estimates, by sectional interests opposed to the reduction of substance abuse. Although few estimates have been made of avoidable costs, those that have been made (see, for example, Collins and Lapsley, 1996 and 2002) indicate that avoidable costs represent in the order of fifty per cent of aggregate costs. In the Australian context this level of avoidable costs provides potentially very high rates of return on expenditures to reduce substance abuse (Collins and Lapsley, 1999 and 2004).

Avoidable cost estimates do not, on their own, indicate the rates of return which the community might achieve. Cost benefit analysis, or at the very least cost effectiveness analysis, is necessary to produce this type of information. However, without knowledge of the avoidable component of substance abuse costs it is difficult to undertake meaningful cost-benefit analysis of prospective expenditures on prevention and/or treatment.

The process of estimating social costs involves estimating the relevant avoidable proportion of each of the cost categories and applying these proportions to each of the relevant aggregate cost estimates. Since not all substance abuse costs fall on the government sector, governments are also likely to show an interest in the proportions of *budgetary costs* which are avoidable. The information developed for estimating the avoidable proportions of aggregate costs can also be readily applied to the estimation of avoidable budgetary costs.

1.3 Reasons for estimating avoidable costs

The original *Guidelines* laid out the four major purposes of guidelines for estimating the aggregate social costs of substance abuse. These were:

1. Economic estimates are frequently used to argue that policies on alcohol, tobacco and other drugs should be given a high priority on the public policy agenda.
2. Cost estimates help to appropriately target specific problems and policies.
3. Economic cost studies help to identify information gaps, research needs and desirable refinements to national statistical reporting systems.
4. The development of improved abuse cost estimates offers the potential to provide baseline measure to determine the efficacy of drug policies and programs.

Avoidable cost estimation can be justified on the same basis.

1.3.1 Priority for substance abuse expenditures

In most countries, the allocation of public funds between competing programs is substantially influenced by pubic servants trained in economics or finance, who are looking to maximize the social rates of return on public expenditures. Accordingly, as part of the decision-making process, they utilize information on economic evaluation of the proposed expenditures. While aggregate cost estimates indicate the economic impact of substance abuse, they do not indicate what proportions of these costs are avoidable, and over what period of time. It is possible that some forms of substance abuse are much more susceptible to prevention measures than others, and so may yield higher gross benefits. Thus, avoidable cost estimates give a better indication of the potential benefits of anti-abuse programs and policies, though they still do not indicate the costs, and the rates of return, of appropriate programs. This issue is considered further below.

1.3.2 Appropriate targeting of specific problems and targets

Estimates of the avoidable costs of substance abuse, in conjunction with estimates of the aggregate costs, provide an extremely valuable information resource for policy analysis and design. This information is of two types:

- **Cost estimation data**. In the process of developing cost estimates, researchers generate virtually all the information necessary to value the benefits (that is, the reduction in social costs) of programs, once their physical outputs (in terms of , say, improved health outcomes, lower work absenteeism and reduced crime rates) have been determined:
- **Information about the physical outcomes potentially available to specific programs**. Essentially, the estimation of avoidable costs involves an examination of the extent to which it is possible to reduce substance abuse and, in the process, what policy measures should be implemented in order to achieve this reduction. As will be seen, this effectively involves an examination of best practice in a range of comparable jurisdictions. Thus the development of avoidable cost estimates will involve an international survey of substance abuse policies and outcomes, and the provision of the resulting information on a systematic basis.

Accordingly, avoidable cost estimation will facilitate both program analysis (through cost benefit analysis) and design of appropriate programs likely to yield the best achievable social rates of return.

1.3.3 Identification of information gaps and research needs

Systematic analysis of avoidable costs is likely to highlight information gaps which, when rectified, will facilitate improved policy design. This information will include answering such questions as:

- What substance abuse outcomes are potentially achievable?
- How best may these outcomes be achieved?
- Over what periods of time would these outcomes be achievable?

Since this information, in terms of both scientific analysis and policy analysis, will inevitably not be static, it will be necessary to review and update information concerning potential best practice on a regular basis.

1.3.4 Provision of baseline measures to determine the efficiency of drug policies and programs

Once policy makers have at their disposal measures of best practice substance abuse outcomes, they will be able much more efficiently to evaluate their existing anti-abuse policies and their allocation of public expenditures to and between anti-abuse programs. There is every reason to believe that the provision of avoidable cost estimates should lead to better policy design.

1.4 Reasons for producing avoidable cost guidelines

Almost no research work has been undertaken internationally to estimate the avoidable costs of substance abuse. There has, therefore, been little development of the necessary underlying theory, and data have not been systematically collected for this purpose. These research difficulties in this area are compounded by the fact that research skills from various disciplines are needed, including criminology, economics and epidemiology.

The objective of guidelines in this area of research is to provide the theoretical and data framework for the estimation of the avoidable costs of substance abuse. Like the original *Guidelines*, this publication is intended to provide a framework for analysis rather than a rigidly prescribed methodology.

It is hoped that these guidelines will encourage and facilitate the development in various countries of avoidable cost studies, leading to the provision of better information bases for the determination of national policies to counter substance abuse. Over time the underlying estimation methodologies will develop and improve. In addition, improvements in scientific knowledge (particularly in the area of epidemiology and criminology) and research methodology will inevitably occur. This suggests that regular revision of these guidelines is indicated.

2 Social costs

2.1 The types of social costs attributable to substance abuse

The aggregate cost *Guidelines* (Single *et al*, 2003) identify the main categories of substance abuse costs. These are summarized in Table 2 below, which also indicates which categories of costs are relevant to which drugs.

Table 2 – Social costs associated with substance abuse

		Alcohol	Tobacco	Illicit drugs
Tangible				
	1. Consequences to health and welfare system	√	√	√
	2. Productivity consequences in the workplace and the home	√	√	√
	3. Crime, law enforcement and criminal justice	√	√	√
	4. Road accidents	√	**X**	√
	5. Fires	√	√	**X**
	6. Environment	√	√	√
	7. Research and prevention	√	√	√
Intangible				
	8. Loss of life	√	√	√
	9. Pain and suffering	√	√	√

Note: the symbol √ indicates relevant and the symbol **X** indicates not relevant

Most substance abuse cost studies provide a greater cost disaggregation than presented in Table 2. The following is a detailed disaggregation of the tangible cost categories. This list is intended to be indicative rather than exhaustive. Where individual costs categories are not comprehensively dealt with in the original *Guidelines*, they are dealt with in more detail later in this report.

Tangible cost categories

1. Consequences to health and welfare system
1.1 Medical
1.2 Hospital
1.3 Nursing homes
1.4 Pharmaceuticals
1.5 Ambulances
1.6 Research and prevention
1.7 Welfare administration
2.1 Productivity consequences in the workplace
2.1.1 Reduction in paid workforce
2.1.2 Absenteeism
2.1.3 Reduced on-the-job productivity
2.2 Productivity consequences in the home
2.2.1 Reduction in unpaid workforce
2.2.2 Sickness
3. Crime, law enforcement and criminal justice
3.1 Law enforcement
3.2 Courts
3.3 Legal charges
3.4 Incarceration and corrections
3.5 Violence
3.6 Property damage
3.7 Lost productivity of prisoners
3.8 Lost productivity of criminals
3.9 Insurance administration
4. Road accidents
4.1 Productivity in the workplace
4.2 Productivity in the home
4.3 Health care
4.4 Law enforcement
4.5 Legal charges
4.6 Incarceration
4.7 Vehicle damage
4.8 Insurance administration
5. Fires
5.1 Productivity in the workplace
5.2 Productivity in the home
5.3 Health
5.4 Fire services
5.5 Property damage
5.6 Insurance administration
6. Environment
6.1 Clean up
6.2 Pollution
7. Research and prevention
7.1 Research
7.2 Prevention

It is probable that the basis for estimating the avoidable percentage of some cost categories may be equally applicable to other cost categories. For example, the avoidable proportion of alcohol-attributable medical costs will probably also be applicable to alcohol-attributable hospital costs. However, for the purposes of systematic research, it is necessary to list all cost categories individually.

2.2 Health impacts of substance abuse

Table 3 lists all the conditions which a recent WHO international study (Ezzati *et* al, 2004) has concluded are causally *and quantifiably* linked to the abuse of alcohol, tobacco or illicit drugs. Quantifiability is very important since, if causal relationships are not quantifiable, it is not possible to estimate the costs of substance abuse or the potential benefits from appropriate anti-abuse policies. In practice, this list of quantifiable drug-attributable diseases has been steadily growing and other studies have produced different lists (see, for example, Appendix B which lists a substantially greater number of conditions which Ridolfo and Stevenson (2001), an Australian study, concluded were causally and quantifiably linked to substance abuse). There are also an increasing number of diseases for which causal relationships with substance abuse have been established but not yet quantified.

Table 3 – Drug-attributable diseases for which the WHO has estimated attributable fractions

Alcohol	GBD	Tobacco	GBD	Illicit drugs	GBD
Low birth weight	50	Tuberculosis	3	HIV/AIDS	9
Mouth and oropharynx cancers	61	Respiratory infections	38	Drug use disorders	90
Oesophagus cancer	62	Nutritional deficiencies	53	Unintentional injuries	149
Liver cancer	65	Mouth and oropharynx cancers	61	Self inflicted injuries	157
Breast cancer	69	Oesophagus cancers	62		
Other neoplasms	78	Stomach cancer	63		
Diabetes mellitus	79	Colon and rectum cancers	64		
Unipolar depressive disorders	82	Liver cancer	65		
Epilepsy	85	Pancreas cancer	66		
Alcohol use disorder	86	Trachea, bronchus and lung cancers	67		
Hypertensive heart disease	106	Melanoma and other skin cancers	68		
Ischaemic heart disease	107	Breast cancer	69		
Cerebrovascular disease	108	Cervix uteri cancer	70		
Cirrhosis of the liver	117	Corpus uteri cancer	71		
Road traffic accidents	150	Ovary cancer	72		
Poisoning	151	Prostate cancer	73		
Falls	152	Bladder cancer	74		
Drowning	154	Lymphomas and multiple myeloma	75		
Other unintentional injuries	155	Leukaemia	76		
Self inflicted injuries	157	Other malignant neoplasms	77		
Violence	158	Other neoplasms	78		
Other intentional injuries	160	Diabetes mellitus	79		
		Endocrine disorders	80		
		Sense organ diseases	98		
		Cardiovascular diseases	104		
		Chronic obstructive pulmonary disease	112		
		Asthma	113		
		Other respiratory diseases	114		
		Peptic ulcer disease	116		
		Appendicitis	118		
		Other digestive diseases	119		
		Genito-urinary diseases	120		
		Skin diseases	124		
		Musculoskeletal diseases	125		
		Oral conditions	143		

Source: Ezzati *et al* (2004), volume 3 (CD-ROM)

GBD signifies Global Burden of Disease code. For the translation from GBD codes to ICD-9 and ICD-10 codes see Mathers *et al* (2003)

3 Avoidable costs: the Feasible Minimum

3.1 Introduction

Where a particular event or medical condition can have more than a single cause it is necessary to have estimates of drug-attributable fractions. For instance, alcohol is not the only cause of road accidents nor is smoking the only cause of fire deaths. If the smoking-attributable fraction for lung cancer were estimated to be 0.84, it would then be known that 84 per cent of lung cancer cases were caused by smoking, the remaining 16 per cent of cases being attributable to other causes. In the absence of the relevant attributable fraction it is impossible to attribute the correct proportion of the total harm to substance abuse. In almost all cases the use of direct measures involves knowledge of attributable fractions.

This requirement can represent, in some areas of harm, a major difficulty in cost estimation. Calculation of attributable fractions (also known as aetiological fractions or attributable proportions) requires two fundamental pieces of information – the relative risk (measuring the causal relationship between exposure to the risk behaviour and the condition being studied) and prevalence (measuring the proportion of the relevant population engaging in the risky activity). For some types of harm the relative risk can be assumed to be similar for genetically and economically similar populations. Applying the estimated prevalence for each population to the relevant relative risk will yield attributable fractions which can be used to estimate harm in various populations (countries).

In practice, the attributable fractions in cost studies are derived from estimates of relative risks derived from research studies across a range of comparable countries. Generally, the calculations of the attributable fractions for all countries have in the past derived from studies conducted in a large set of countries with similarly high levels of economic development.

However, for some types of harm it would be quite unsafe to assume similar relative risks even in countries at similar levels of economic development. The most obvious type of harm in this context is crime where, for a range of cultural, social, legal and other reasons, relative risks can vary greatly even between countries with similarly high *per capita* incomes. A good example of the different rates of this type of harm is alcohol-attributable violence where experience varies greatly, even among Western European countries.

The manner in which attributable fractions are derived allows us to estimate how many deaths and hospitalizations are attributable to substance abuse. If there had been no substance abuse, then all of these deaths and hospitalizations would not have occurred. But that does not mean that all of those deaths and hospitalizations were avoidable. There are at least three major difficulties with considering all deaths and hospitalizations caused by substance abuse to be avoidable:

1. Even if substance abuse were to end immediately, many deaths and hospitalizations (and other adverse consequences which lead to costs to the economy) would continue due to the lagged effects of past substance abuse.
2. When a risk factor (like substance abuse) causes a death or hospitalization, in a sense it may prevent another risk factor from causing a death or hospitalization. The obvious example is the fact that mortality is inevitable: everyone dies eventually from something or another, and if it is not due to substance abuse, it will be due to another cause.
3. There would necessarily be huge policy costs to end all substance abuse (in the unlikely event that this were possible), and it virtually impossible to estimate these costs as such policies have not even been identified, let alone costed.

The extent of the first problem is very difficult (and in practical terms, virtually impossible) to estimate. Given a set of assumptions concerning time intervals between different levels of use or risky behaviours and the onset of disease or death (information that is generally lacking), one could attempt to make estimates of lagged effects for some causes. But even if this could be done (which is highly doubtful), it would be limited to only a small subset of diagnoses. And in any case, the exercise would be of little use due to the second and third problems. Even with perfect information on the aetiology of substance-related diseases and accidents, researchers must still confront the other two problems.

It could be considered that the term substance abuse is misleading, because a significant portion of the burden attributable to drugs is caused by use only, i.e. by individuals who do not fall under the gold standard definitions of substance dependence or abuse (Rehm, 2003). For example, a road accident may be attributable to the intoxication (as defined by legislated maximum permissible blood alcohol content) of a driver who, nevertheless, would not fulfil any criterion of dependence or abuse in the International Classification of Diseases, Tenth Revision (ICD-10) (WHO, 1992-1994) or the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV-R) (American Psychiatric Association, 1994). Such alcohol use is often labelled as alcohol misuse. On the other hand, much alcohol use can, by any definition, clearly be categorised as abuse

Since social cost studies are essentially economic studies, they require a definition of substance abuse which is meaningful in economic terms. To economists, substance abuse exists when substance use involves the imposition of social costs additional to the resource costs of the provision of that drug. Abuse occurs if society, including the substance user, incurs extra costs as a result of the drug use. Any use of illicit substances is deemed to be abuse since if use is deemed to be illegal it is clearly considered by society to be abuse

For the purposes of simplicity of expression in these guidelines the term substance abuse is used throughout and is defined as consisting of tobacco abuse, illicit drug abuse, and alcohol abuse and misuse.

3.2 The concept of avoidability

The social costs of substance abuse worldwide are high (see, for example, Single *et al*, 1998; Harwood *et al*, 1998; Collins and Lapsley, 2002; and Andlin-Sobocki and Rehm, 2005). These costs are mainly related to the costs of health care, crime and law enforcement, and losses in productivity. The substance abuse-related burden of disease is composed of the mortality and morbidity attributable to the abuse of alcohol, tobacco, and illegal drugs and is one of the main underlying components in substance abuse-related social costs. Knowing the overall attributable costs of substance abuse or the overall substance-abuse attributable disease burden has been found unsatisfactory, as it is not clear which proportion of the costs or the burden could in principle be changed. This changeable part has been labelled the avoidable cost or avoidable burden (WHO, 2002). Estimating avoidable burden is an important element in the process of estimating avoidable costs.

The first step in estimating avoidable burden is to conceptualise the attributable burden of disease; that is, the burden of a given disease in a given population that is identified as due to a specific exposure to a risk factor or multiple risk factors. Consequently, that portion of disease burden could, *in principle*, be reduced or eliminated if the causative exposure is reduced or eliminated. Attributable burden is conceptualized regardless of whether such a reduction is achievable in practice or not.

Based on the conceptualization of attributable burden, it is then possible to introduce the term avoidable burden of disease. The latter term denotes the proportion of disease burden that can be reduced by changing the current exposure distribution to an alternative, more favoured, exposure distribution. Clearly, the size of the avoidable burden caused by a given risk factor will always be smaller than or, at most, equal to the burden attributable to that risk factor. Little has been written on the problems of estimating avoidable burdens/costs and so a discussion is presented below of a method to estimate avoidable burden specifically for substance use as a risk factor, with special emphasis on methodological problems and potential solutions.

3.3 Avoidability, optimality and zero tolerance

The *avoidable* burden/costs discussed here should be contrasted with the economist's concept of the *optimal* level of substance abuse. Economists argue that the optimal level of drug consumption is reached when the incremental cost to the community as a whole of achieving a given reduction in consumption is exactly matched by the incremental benefit to the community of that reduction. If the incremental benefit is greater than the incremental cost, achieving optimality would require a further reduction in consumption. If the cost exceeds the benefit, then consumption has been reduced to sub-optimal levels.

The concepts of avoidability and optimality can lead to quite different outcomes. It is perfectly possible that optimal levels of consumption may not be achievable. For example, optimal levels of tobacco consumption may be well below the levels which are achievable in a real world in which severe constraints exist on the public resources available to achieve reductions in smoking prevalence. On the other hand, it may be technically possible to reduce tobacco consumption below levels which economists judge

to be optimal. In any event, there are likely in practice to be severe informational problems in determining optimal consumption levels. These guidelines concentrate on issues relating to the estimation of the *avoidable costs* of substance abuse, not the *optimal levels* of substance abuse.

Some public health advocates consider that the target for public health interventions should be a zero level of substance abuse. This might be called a zero tolerance approach. There are, for the purposes of these guidelines, problems with this approach from an economic perspective.

Economists would argue, as explained above, that in most situations the optimal outcome is not zero substance abuse but a level at which the additional costs of reducing abuse further are matched by the additional social benefits of that reduction. Even if zero substance abuse were achievable, in most instances the costs of achieving that outcome would exceed the benefits. In other words, the resources could be used more productively elsewhere.

In practice, zero abuse is not likely to be an achievable outcome. The concept of avoidable costs relates to what is achievable in the real world, not what would be desirable in an ideal world of unlimited resources.

3.4 Approaches to calculation of the Feasible Minimum

From an economic policy perspective, it is necessary to determine the maximum quantifiable, measurable reduction in substance abuse costs which effective policies can be expected to achieve. The lowest achievable level of substance abuse can be termed the Feasible Minimum, and the four methodologies discussed here can be regarded as ways in which a Feasible Minimum can be calculated in order to provide policy objectives. Of course, such avoidable cost estimates assume that the calculation of total cost estimates will have already been undertaken. Various possible approaches exist for estimation of the Feasible Minimum.

One method of achieving an estimate of a Feasible Minimum is the use of the classic epidemiological approach, deriving the attributable burden from calculations of relative risk and the prevalence. From this calculation of the attributable burden, both past and future risk factor distributions can be estimated, which also provide data to enable the calculation of a Feasible Minimum. This approach can be modelled to demonstrate the difference between the attributable and the avoidable burden.

Armstrong applied this concept in his work on prevention, using a measure which he described as the “Arcadian normal” as his feasible minimum for preventable mortality for a range of conditions. Instead of using epidemiological data from which to calculate the feasible minimum, he used as a comparator the lowest recorded rate (e.g. of lung cancer) which had been achieved by a country which could be considered a reasonable comparison with the study country. From that comparator, the amount which can be prevented can be estimated, and from that the calculation for avoidable costs can be made.

This concept does not in itself suggest that similar policies, regulations, and even health behaviours are necessarily appropriate nor transferable from the comparator country, but simply indicates what amount of burden reduction and costs have been able to be achieved. A disaggregation of effective policies from a comparator country may be useful in developing economic evaluation studies, and provide guidance on resource distribution which could be made to different policy implementations.

Another possible approach to estimation of the Feasible Minimum is to use recently-published WHO data on drug-attributable fractions, morbidity and mortality. These data can be used to identify best performance among countries in sub-regions identified by the WHO as having common characteristics. This approach could be adopted for countries where insufficient domestic data are available. Although the only practical solution for many countries, it should be noted that this method is likely to be highly imprecise, given the multiple layering of assumption underlying the application of attributable fractions from one setting to another.

There may also be circumstances in which evidence about the known effectiveness of specific interventions may be used in avoidable proportions. This constitutes a fourth type of approach.

All four approaches are described in more detail below. Inevitably there are difficulties and complexities both in deriving the attributable fraction and then in calculating the proportion which is avoidable, which indicates the Feasible Minimum.

3.5 From the classical to the distributional approach in epidemiology

Two different approaches to the epidemiology of risk factor attribution can be broadly distinguished: the classical epidemiological approach where, on the basis of a 2x2 table, disease risk can be estimated with respect to exposure and non-exposure of various populations to a single specific risk factor. Based on this categorically derived relative risk and the prevalence, the attributable burden can be derived, resulting in an estimate for a population who have *already* been exposed, i.e. focussing on the past. The main counterfactual scenario (Maldonado and Greenland, 2002) in this approach asks the question: “What would have happened if no exposure had occurred?”

More modern developments of epidemiology ask the question: “What would happen if risk factor distributions shifted to different counterfactual scenarios?” (Murray and Lopez, 1999). The modern approach not only looks at distributional shifts at one time, but also takes the future time dimension into consideration, and thus is able to predict future developments.

The current contribution focuses on the second approach. It thus conceptualizes the avoidable burden by specifying alternative scenarios for risk factor distributions, including potential future distributions. It is important to look at the impact of a risk factor over time, not just an epidemiological snapshot of today’s attributable burden, in order to fully estimate the burden of disease contributions of acute and chronic disease burden which might be avoided in the future. This is an especially important factor for

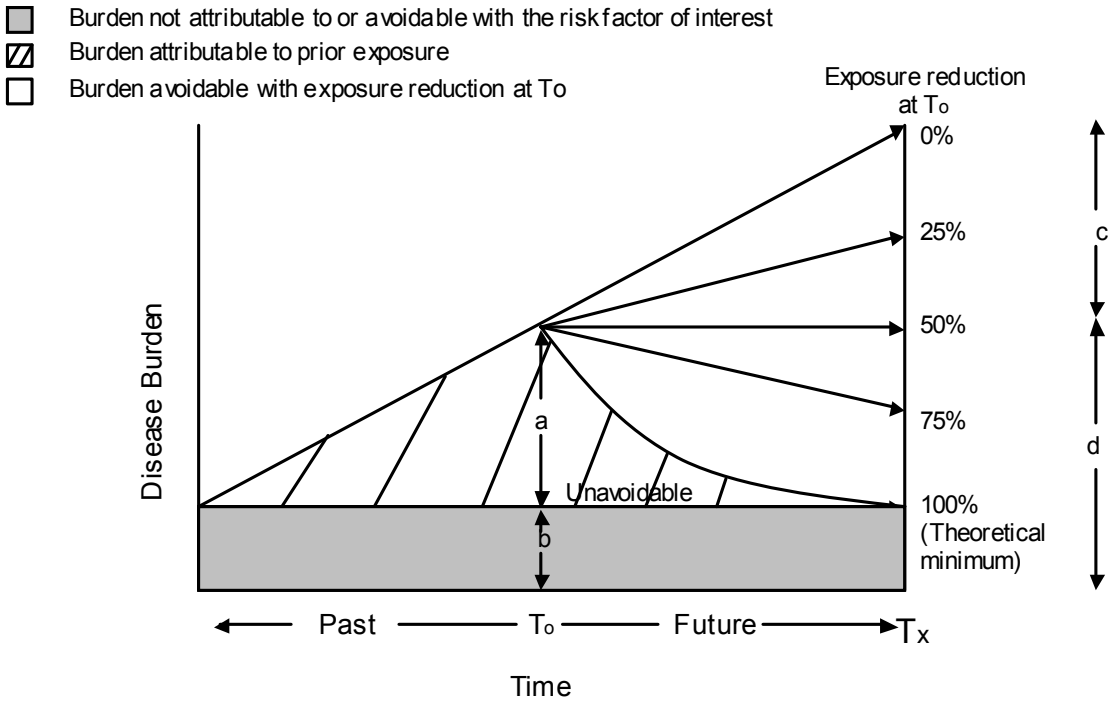
policy. We will also consider potential interactions and competing risks between influencing factors.

Note that the proposed framework is conceptualized independently from the measure for burden of disease used, whether it be mortality, morbidity, or any summary measure (such as the traditionally used Disability Adjusted Life-Year (DALY) (Murray and Lopez, 1997)), so that it can be universally applied.

3.5.1 An epidemiological model to estimate avoidable burden

Figure 1 illustrates the conceptual model of the difference between attributable and avoidable burden, cited from the epidemiological model of Murray *et al* (2003).

Figure 1 – A conceptual model of attributable and avoidable risk with increasing projected burden



a Disease burden at T_o attributable to prior exposure

b Disease burden at T_o not attributable to the risk factor of interest (caused by other factors only).

The burden not attributable to the risk factor of interest (grey area) may be decreasing, constant or increasing over time. The constant case is shown in the figure.

c Disease burden avoidable at T_x with a 50% exposure reduction at T_o.

d Remaining disease burden at T_x after a 50% reduction in risk factor exposure

Source: Murray, C.J.; Ezzati, M.; Lopez, A.D.; Rodgers, A.; and VanderHoorn, S (2003). "Comparative quantification of health risks conceptual framework and methodological issues", *Population Health Metrics*, 1(1): 1-20.

Consider the disease burden at time T_0 (that portion which is attributable to prior exposure), denoted by the letter *a* in Figure 1. This is all the burden of disease which can be attributed to prior exposure of the risk factor under consideration before T_0.

In the example, a general situation is given where the background burden, i.e. the burden due to other factors except the risk factor under consideration, is constant over time. This background burden is denoted by the letter *b* at time T_0 and, because it is constant over time in this specific example, the burden is the same size at all time points. Of course, in other situations the burden attributable to other factors than the risk factor under consideration may fluctuate.

The burden attributable to the risk factor under consideration in the example is increasing constantly over time until time T_0. Then different scenarios are shown. Let us discuss three of them. If nothing (e.g. intervention, changes in cultural acceptance) happens at time T_0, the attributable burden continues to increase linearly. On the other hand, if the exposure is reduced completely, then the attributable burden is decreasing until time T_x, when it reaches zero.

Consider tobacco use as an example in a society where prevalence rates have been increasing and would continue to increase without any intervention. If some drastic intervention could reduce smoking completely at a certain time point, smoking-related disease burden would not be zero the moment thereafter. Instead, some burden of disease would persist, e.g. burden of disease due to existing tobacco-related lung cancer, and some people may even develop new lung cancer based on their past exposure.

Now consider a reduction of exposure to the risk factor by 50 per cent at time T_0. In the example, such a reduction would mean that a constant attributable burden would result. This burden is, of course, a mixture of the impact of past exposure (i.e. prior to T_0) plus the impact of the exposure after T_0.

Three main components have to be known to estimate this model. They are:

1. The relationship between exposure (i.e. the risk factor under consideration) and disease burden (attributable risk) and its temporal trend
2. The amount of burden caused by other factors than the risk factor under consideration and the temporal stability of this burden.
3. The specific trajectory of burden reduction after exposure reduction

Methods to estimate the relationship between a certain exposure distribution and disease burden are well established (e.g., odds ratios, relative risks; see Rothman and Greenland, 1998). Similarly, methods to estimate the proportion of burden attributable to the distribution of a certain risk factor were developed some decades ago and were first described by Miettinen (1972) in the early 1970s and Walter a little while later (Walter, 1976; 1980). They have since been used to generate estimates of the attributable burden for substance use and other risk factors around the world (Ezzati and Lopez, 2000). The key concept used herewith is that of an attributable fraction (AF; also called the aetiologic fraction). It depends on the exposure distribution (or, in the discrete case, on the prevalence of different exposure categories) and on the relative risk for burden related to the respective exposure levels.

Below are the two formulas for the case of continuous exposure levels and the discrete case, both of which are fully developed and described elsewhere (Walter, 1976; 1980; Eide and Heuch, 2001; Murray *et al*., 2003).

The contribution of a risk factor to disease can be estimated by comparing the burden due to the observed exposure distribution in a population with that from another *distribution* (rather than a single reference level such as non-exposed) as described by the generalized equation shown as Equation 1.

Equation 1

$$PIF = \frac{\int_{x=0}^{m} RR(x)P(x)dx - \int_{x=0}^{m} RR(x)P'(x)dx}{\int_{x=0}^{m} RR(x)P(x)dx}$$

where *PIF* is the "potential impact fraction", a generalized form of the attributable fraction; $RR(x)$ is the relative risk at exposure level x, $P(x)$ is the population distribution of exposure, $P'(x)$ is the counterfactual distribution of exposure, and m the maximum exposure level. As Murray et al. (2003) have noted, this formula can be further generalized to deal with a situation, where the relative risks change in the counterfactual scenario.

The corresponding relationship when exposure is described as a discrete variable with k levels is given by Equation 2.

Equation 2:

$$PIF = \frac{\sum_{i=1}^{n} P_i \ RR_i - \sum_{i=1}^{n} P_i' \ RR_i}{\sum_{i=1}^{n} P_i \ RR_i}$$

where
i: exposure level category
RR_i: relative risk at exposure level i
P_i: prevalence of the i*th* category of exposure

The concept of attributable fraction (or the generalized form of PIF), as defined here, can only describe a snapshot at a specific time. Attributable fractions without including a time dimension are not able to characterize those cases whose occurrence would have been delayed or in part prevented due to exposure reduction (see Greenland and Robins, 1988). As a remedy, Greenland and Robins (1998) recommend the use of aetiologic fractions with a time dimension to account for this shortcoming. Time-based measures are discussed in some detail by Murray *et al* (2003) in their seminal conceptual framework of the comparative quantification of health risks.

The theoretical minimum risk (see Figure 1) is trickier to define than the exposure-burden relationship. Theoretical minimum risk denotes the exposure distribution that would result in the lowest population burden, irrespective of whether such a distribution is currently attainable in practice (Murray and Lopez, 1999). In the example it was set at zero attributable burden, but this is not necessarily always the case. Consider alcohol, and assume only one relevant exposure dimension in disease aetiology (see below for a discussion of volume and patterns of drinking as a two-dimensional exposure association), such as average volume of drinking. Minimal risk then occurs at zero (i.e. no drinking at all) for most related diseases such as cancer or traffic injury, but not for some whose risk actually decreases at exposures greater than zero e.g. heart disease (Rehm *et al*, 2003a; Rehm *et al*, 2003b; Rehm *et al*, 2004). For composite outcome measures, e.g. all-cause mortality, there is also reason to believe that the level of the exposure associated with minimum burden is greater than zero (i.e. some level of drinking) (Rehm *et al*, 2001; Gmel *et al*, 2003). The exact value of exposure associated with the minimal burden will depend on the composite measure used, and the disease distribution in the country or region examined. Thus, the theoretical minimum risk will fluctuate across cultures.

A theoretical minimum risk with an exposure greater than zero has interesting implications. For alcohol in the above example it means that, even if drinking occurs at the theoretical minimum, there will be some attributable disease burden. For instance, in a society with relatively large coronary burden and assuming a theoretical minimum risk for the population occurring at approximately 1 drink/day, there will be disease burden associated with such an exposure level of moderate drinking, e.g., for certain gastrointestinal diseases (Taylor *et al*, 2005) or accidents (Rehm and Gmel, 2003).

The above example is hypothetical for several reasons:

a) It assumes all people having the same exposure level (1 drink/day for all people) rather than a distribution of exposure, known to exist for all populations drinking alcohol. (Skog, 1985). Note that the above statement only refers to the existence of a distribution and does not specify the exact shape of it (e.g., lognormal vs. Poisson).

b) It completely disregards the fact that alcohol has at least two dimensions relevant for disease, average volume of alcohol consumption and patterns of drinking (Rehm *et al*, 2003a).

c) There is no known intervention, which would result in the above-mentioned distribution of 1 drink/day for all (Babor *et al*., 2003; Room *et al*., 2005).

In more general terms, the following points can be made:

a) We should always model shifts in risk factor distributions when estimating avoidable burden of disease.

b) We should always model all relevant dimensions of a risk factor when estimating avoidable burden of disease.

c) We need other counterfactual scenarios in addition to the theoretical minimum risk, when estimating avoidable burden of disease. The question will be where to obtain the counterfactual scenarios from.

The trajectory of the burden reduction after changes in exposure is also difficult to define, not only because it should incorporate changes of exposure reduction as well, but because it also has to make estimates for the change of relative risk over time of several disease outcomes, potentially both acute (e.g., alcohol and deaths from drinking and driving) and chronic (e.g. alcohol and chronic pancreatitis). For instance, after a change of smoking status to abstinence, we have to know the relative risk of an ex-smoker after one year after abstaining, two years after abstaining, three year after abstaining, and so on. For acute outcomes, the problem is much easier. Once the prevalence of alcohol, tobacco or illicit drugs is reduced, all the acute outcomes (e.g. injuries) are reduced accordingly. To give an example: while drinking over the past years does affect the cancer risk of people today, even if they started abstaining in between, it does not affect the traffic accident risk.

3.5.2 An example: shifting the smoking distribution in Canada by 10 per cent

Below is a real example of applying distributional shifts in exposure to tobacco consumption in Canada using smoking prevalence data from the 2003 *Canadian Community Health Survey* and mortality data from the Statistics Canada in 2002. A hypothetical decrease in exposure of 10 per cent was run, i.e. the prevalence of each smoking category was reduced by 10 per cent shifting into the next lower category, and the 10 per cent from the lowest smoking category shifted into the former smoker category. To give an example: for men in the second highest smoking category, i.e. smoking between 15 and 24 cigarettes/day. The prevalence in this category was 6.55 per cent before the shift. Then, 10 per cent of the of 6.55 per cent or 0.66 per cent went to the next lower smoker category, i.e. smoking between 1 and 14 cigarettes/day, but 0.55 per cent were added coming from the highest smoking category, resulting in a net decrease of 0.11 per cent in this category. All other cells were modelled accordingly. The resulting changes in prevalence are shown in Table 4.

Table 4 – Resulting prevalence of tobacco smoking in per cent after a -10 per cent shift in exposure distribution in Canada 2003, by gender

	Men		Women	
Smoking Categories	**Before**	**After**	**Before**	**After**
Never	29.12	29.12	40.39	40.39
Former	44.59	45.18	37.80	38.29
Current/Occasional	5.90	6.14	4.90	5.34
1-14 cigs/ day	8.35	8.17	9.28	8.80
15-24 cigs/ day	6.55	6.44	4.51	4.37
25+ cigs/ day	5.49	4.94	3.12	2.81
Total	**100.00**	**100.00**	**100.00**	**100.00**

Table 4 shows the gender-specific shift in prevalence of smoking for different categories before and after a 10 per cent decrease in each exposure prevalence category as described above. Overall, smoking prevalence decreased by 2.2 per cent on both genders. The impact this shift distribution had on tobacco-related disease-specific mortality was modelled based on the mortality data for 2002. The baseline scenario was taken from the second Canadian study on social costs of substance abuse where tobacco-related attributable burden was modelled gender, age and disease specific with a counterfactual scenario of zero smoking as the level denoting minimal risk (Baliunas et al., 2005). To estimate avoidable burden, the shift in prevalence was again modelled specifically by gender, age and disease, resulting in the reduction of smoking-related mortality summarized in Table 5. The relative risk to denote the relationships between category of exposure and outcome were taken from meta-analyses (for details see Baliunas et al., 2005). More than 20 different disease categories had to be modelled in order to arrive at Table 5.

Table 5 – Impact of exposure shift of -10 per cent in exposure to tobacco smoking on disease-specific, tobacco-attributable mortality in Canada, 2002 (number of deaths)

	Lung Cancer		Other Cancers		Cardiovascular Disease		Respiratory Disease		Passive Smoking (lung cancer and IHD)		Total[†]	
	M	**F**	**M**	**F**	**M**	**F**	**M**	**F**	**M**	**F**	**M**	**F**
Before Shift	9127	4753	4199	1795	8476	6604	4776	3810	757	528	27563	17655
After Shift	9026	4553	4160	1784	8267	6394	4761	3797	683	475	27117	17164
Diff	101	200	38	12	209	210	15	13	74	53	447	491
Total	**301**		**50**		**419**		**28**		**127**		**938**	

[†]Totals represent mortality due to all outcomes, not row totals of those selected

A distributional shift of 10 per cent in the exposure level to tobacco smoke resulted in 938 fewer tobacco-attributable deaths, 2.1 per cent less than before the shift. The largest absolute differences before and after were seen for lung cancer and cardiovascular diseases, and also for deaths due to lung cancer and ischemic heart disease attributable to passive smoking. However, relative differences were most pronounced for the acute outcomes.

The above example illustrates the possibilities of modelling avoidable burden using the above specified framework based. It also illustrates which points still have to be improved. First, smoking-related attributable burden is still modelled in a way using current levels of smoking as an indicator of cumulated past exposure. While this procedure is usual in the literature and the basis of the most-used software for calculation smoking-attributable mortality, it clearly introduces measurement error, which may increase, as smoking behaviour across the lifespan does not seem to be as stable as it used to be. Second, the distributional shift is somewhat arbitrary in two ways: it is not clear why a 10 per cent shift should be the basis for avoidable burden, nor are the actual shifts in prevalence empirically based. For instance, in interventions underlying reductions in smoking prevalence such as taxation, different shifts in smoking prevalence distributions might be seen (e.g. changing from the highest smoking level into abstention). Finally, it is again a snapshot picture, depicting one change without incorporating the time dimension.

However, to generally select which shifts should be modelled, and how realistic they are, it may still be helpful to look at similar countries or historical trends as comparators. Thus, one would first screen plausible developments based on historical trends or the distribution in other countries, and then model shifts in risk factor distributions and subsequently the avoidable burden associated with this shift. For the Canadian example below, one would look at the distributional changes resulting from intervention packages being used in countries or regions similar to Canada and model avoidable burden accordingly.

3.6 The Arcadian normal

A second type of approach to estimating the Feasible Minimum is by estimating what has become known as the *Arcadian normal*. Pioneering work in this area was done by Armstrong (1990). His work was expressed in terms of preventable mortality and morbidity but it is reasonable to extend the concepts embodied in his work to other costs such as the property costs resulting from drug-attributable crime or smoking-attributable fires. He talks of assuming "the existence of some level of disease that might reasonably be achieved if only we knew all that might reasonably be known about the causes of the disease in question and could apply them in practical programmes in the community. There is no simple way of identifying this level of disease but it may be assumed to be less than the lowest level of disease that obtains in some group of genetically similar populations". Armstrong terms this level of disease the 'Arcadian normal' "because it represents the nearest approach that we can make to harmony between humankind and its environment. Arcadia, in ancient Greece, was a region renowned for the contented pastoral simplicity of its people".

Armstrong's approach is to compare the most recently available age standardised mortality rates for a range of causes in a group of countries with genetically similar populations and with similar living standards. He takes the Arcadian normal to be the lowest age-standardised mortality rate for each cause of death in the 20 European countries he studied and from these he estimates the proportions of potentially preventable mortality in Australia. His results are presented in Table 6 below.

Table 6 – Estimates of potentially preventable mortality in Australia

Cause of death	Rate in Australia*	Country with lowest rate	Rate in that country*	Per cent preventable in Australia
All causes	**838**	**Switzerland**	**726**	**13.4**
Infectious and parasitic diseases	**4.3**	**Austria**	**4.0**	**7.0**
All cancers	**197**	**Greece**	**161**	**18.3**
Stomach cancer	10.1	USA	6.0	40.6
Lung cancer	41.0	France	22.2	44.1
Breast cancer	21.2	Spain	19.0	10.4
Circulatory diseases	**410**	**France**	**265**	**35.4**
Ischaemic heart disease	231	France	76	67.1
Cerebrovascular disease	95.6	Canada	57.5	39.8
Respiratory diseases	**64.7**	**Austria**	**42.5**	**34.3**
Chronic bronchitis, emphysema and asthma	16.9	USA	8.3	50.9
Digestive diseases	**29.0**	**Sweden**	**21.1**	**27.2**
Chronic liver disease and cirrhosis	8.7	Ireland	3.5	59.8
Injury and poisoning	**50.4**	**England and Wales**	**34.3**	**31.9**
Road crashes	17.9	England and Wales	8.8	50.8
Suicide	11.8	Greece	3.9	66.9

* Age-standardised mortality rate per 100,000 of the population
Source: Armstrong (1990)

Note that Armstrong's table, as presented in his 1990 paper, is reproduced above solely to illustrate the concept of the Arcadian normal. It is not suggested that the actual estimates of the normals presented in that paper should be used in future studies. Clearly new information has become available in the years since Armstrong's important paper was published.

If it is limited to historical examples, the Arcadian normal procedure can fail to make appropriate adjustment for long-term global trends such as the trend away from smoking. Smoking rates are declining significantly across the board in many regions (e.g., North

America, Western Europe, Oceania). Even though the Arcadian normal expressed in terms of smoking prevalence may be 15 per cent now, it will probably be much lower in another decade. This calls to question the whole idea of using historical examples as an absolute standard. Fifty years from now the Arcadian normal for smoking may only be 5 per cent or 10 per cent. Two hundred years ago it would have been close to zero. Arcadian normals are a reasonable approximation for a given place and time period only.

Table 6 above indicates that the more disaggregated is the information on mortality and morbidity, the more likely it is to lead to accurate estimates of the normal. As an illustration, Table 6 shows that the derivation of individual Arcadian normals for specific types of cancer would be more accurate in producing percentages preventable than simply applying the across-the-board figure for all cancers. This is especially true when dealing with drug attributable diseases, since the drug attribution factors for different diseases will vary widely (and will, in many cases, be zero).

In principle it would be desirable to have individual Arcadian normals for every drug-attributable condition, although this is, in practice, unlikely to be achievable.

In estimating the Arcadian normal, direct or indirect (proxy) measures can be used. Direct measures are physical or financial measures which relate directly to the costs attributable to substance abuse. For example, alcohol-attributable road accidents or smoking-attributable fire deaths are direct measures of harm resulting from substance abuse. Economists can then translate these physical measures into financial costs. However, for almost every harm linked to substance abuse there are multiple causes. Alcohol is not the only cause of road accidents nor is smoking the only cause of fire deaths. Where a particular event or medical condition can have more than a single cause it is necessary to have estimates of substance-attributable fractions. For example, if the smoking-attributable fraction for lung cancer were estimated to be 0.84, it would then be known that 84 per cent of lung cancer cases were caused by smoking, the remaining 16 per cent of cases being attributable to other causes. In the absence of the relevant attributable fraction it will be impossible to attribute the correct proportion of the total harm to substance abuse. In almost all cases the use of direct measures involves knowledge of attributable fractions.

This requirement can represent, in some areas of harm, a major obstacle to the use of direct measures. As indicated above, calculation of attributable fractions requires two fundamental pieces of information – the relative risk and prevalence. For some types of harm the relative risk can be assumed to be similar for genetically and economically similar populations. Applying the estimated prevalence for each population to the relevant relative risk will yield attributable fractions which can be used to estimate harm in the various populations (countries).

Nevertheless, there exist some significant issues with the Arcadian normal which must be acknowledged when making these types of calculations

- First, it is based on disease and thus is not *exposure-specific*. Different countries will still have different, less than optimum, exposure profiles that could be changed to lower the specific disease burden, but the Arcadian normal does not recognize these exposure-specific differences. It just covers broad summary outcome measures.

- Second, the theoretical minimum needed in our framework is for burden of disease across different diseases, not for an individual disease category. Summing up across disease categories has the additional problem that the Arcadian normal will stem from different countries (in Armstrong's analysis, France for CHD, Austria for infectious disease, Sweden for digestive disease etc.) which means different burdens from factors other than the exposure under consideration have to be taken into account.
- Third, countries may be so different with respect to other factors that they may not be comparable as regards achievable disease reduction. There might also be other reasons, such as cultural differences (e.g. Muslim countries with very low alcohol consumption), which explain different exposure levels, and such "feasible" exposure levels might not be really feasibly achieved in other countries.

An additional issue is that the Arcadian normal procedure should require appropriate adjustment for long-term global trends, such as the trend away from smoking. Smoking rates are declining significantly across the board in many regions (e.g., North America, Western Europe, Oceania). Even though the Arcadian normal expressed in terms of smoking prevalence may be 15 per cent now, it will probably be much lower in another decade. This calls to question the whole idea of using historical examples as an absolute standard. Fifty years from now the Arcadian normal for smoking may only be 5 per cent or 10 per cent. Two hundred years ago it would have been close to zero. Arcadian normals are a reasonable approximation for a given place and time period only.

Armstrong's Arcadian normal will not work for the model proposed above because this model needs instead an *exposure-specific* feasible minimum of disease burden. The concept of the Arcadian normal could, however, be transferred to exposure and could define feasible exposure changes based on the minimum of exposure distributions in similar societies achieved by intervention. It would thus introduce a Feasible Minimum (Murray and Lopez, 1999) – defined as an exposure distribution that has already been achieved in comparable societies. This possibility is developed in Section 3.7 below.

Using a Feasible Minimum will result in more practically achievable solutions and thus less avoidable burden than if using a purely theoretical minimum. It would also be possible to estimate pathways towards the avoidable burden in hypothetical scenarios based on changes to current risk factor exposure, making this a powerful statistical tool for policy and intervention development. However, we would still have to deal with problems of comparability between countries.

3.7 Exposure-based comparators

The approaches discussed above impose substantial data requirements, which will be difficult to satisfy in many countries, particularly in developing nations. In these circumstances it may be necessary to resort to less sophisticated but more practical measures. An alternative approach could use prevalence data as a proxy for attributable fractions in estimating the avoidable proportions of the costs associated with drug-attributable morbidity and mortality.

The recent publication by the *World Health Organisation* of *Comparative Quantification of Health Risks* (Ezzati *et al*, 2004) has produced a significant improvement in the availability of information on the relationships between, on the one hand, substance abuse and, on the other, mortality and morbidity. The *Annex* to Ezzati *et al* (*op cit)* provides population attributable fractions for a wide range of substance abuse-disease relationships, classified by age, sex and sub-region. Also provided, cross-classified according to the same variables, are attributable fractions for years of life lost to premature mortality (YLL) and overall disease burden as measured by disability adjusted life years (DALYs). Fuller details of this information are provided in Appendix A below.

As indicated above, the calculation of attributable fractions requires information on both relative risk and prevalence. On the assumption that relative risk is the same for all countries within a WHO-defined sub-region, variations in prevalence within the sub-region could be considered as proxies for variations in attributable fractions (for definitions of the sub-regions see Appendix A). The higher the prevalence rate, the higher will be the attributable fraction. By comparing the prevalence rate in the country under study with the lowest prevalence rate of all countries in the sub-region, an estimate can be made of the percentage of burden, and therefore of costs, that is avoidable. This approach, though simplified, is still consistent with our recommended concentration on *exposure* as an indication of the Arcadian normal, rather than *outcomes*.

A considerable amount of information on the international prevalence of drug use is available from WHO sources, particularly the *WHO Statistical Information Service (WHOSIS)* and the *Organisation's Global Alcohol Database, Global Information System of Tobacco Control* and *Tobacco Atlas* (Mackay and Eriksen, 2002).

Consider the example of smoking prevalence in WHO sub-region AMR-D, defined in Ezzati *et al* (2004) as consisting of Bolivia, Ecuador, Guatemala, Haiti, Nicaragua and Peru. Table 7 below presents adult smoking percentages, classified by sex, for the countries in this sub-region.

Table 7 – Adult smoking prevalence, WHO sub-region AMR-D

Country	**Prevalence rate**		
	Male (per cent)	**Female (per cent)**	**Total (per cent)**
Bolivia	42.7	18.1	30.4
Ecuador	45.5	17.4	31.5
Guatemala	37.8	17.7	27.8
Haiti	10.7	8.6	9.7
Nicaragua	n.a.	n.a.	n.a.
Peru	41.5	15.7	28.6

Note: n.a. indicates not available

Source: Mackay and Eriksen (2002)

In sub-region AMR-D, the country with the lowest adult smoking rates is Haiti. Comparing the prevalence rates of other countries with those of Haiti provides an indication of the proportions of the smoking burdens in these countries which are potentially avoidable. The results of such a calculation is presented in Table 8 below.

Table 8 – Avoidable proportions of smoking burdens, WHO sub-region AMR-D

Country	Male (per cent)	Female (percent	Total (per cent)
Bolivia	74.9	52.5	68.1
Ecuador	76.5	50.6	69.2
Guatemala	71.7	51.4	65.1
Haiti	0.0	0.0	0.0
Nicaragua	n.a.	n.a.	n.a.
Peru	74.2	45.2	66.1

Note: n.a. indicates not available

To take the best performance of the countries in the sub-region as the Feasible Minimum is a more conservative approach than that embodied in Armstrong's original definition of the Arcadian normal (which translates in the approach adopted here to a level *less than* the lowest level of exposure in some group of genetically similar populations). In the example above, the exposure-based comparator would imply that aggregate smoking costs in Haiti cannot be further reduced. A limitation of this methodology is that it suggests that no further reduction in costs can be achieved by the best-performing country. In practice, it will probably be the case that the calculated Feasible Minimum changes as improvements in public anti-abuse policies lead to further declines in prevalence in the country currently having the lowest sub-regional prevalence rates.

The publication of the epidemiological information in Ezzati *et al* means that one of the major obstacles, particularly in developing countries, to the estimation of the aggregate social costs of substance abuse has been removed. Accordingly, it will also become less difficult to estimate the *avoidable* costs of substance abuse in many countries.

It may also be possible to estimate minimum feasible exposure levels to the harmful substances. For example Warner and Burns (2003) have attempted to estimate minimum feasible smoking prevalence levels in the United States. Such estimates may then be utilized in estimating the avoidable proportions of the social costs of substance abuse.

Another approach to estimating minimum levels of prevalence rates might be to use smoking prevalence rates among doctors, on the grounds that they, of any group in the community, would have the best knowledge of the causes and effects of tobacco-attributable diseases. For example, the estimated smoking prevalence rate in Finland in 2001 for male physicians was seven per cent, and for female physicians four per cent (Barengo *et al*, 2004). The estimated smoking prevalence rate for general practitioners in Australia in 1996 was 3.2 per cent (Young and Ward, 1997).

3.8 Using evidence on the effectiveness of interventions

In estimating avoidable costs it may at times be useful also to utilise research evidence on the effectiveness of interventions designed to reduce or alleviate the effects of substance abuse. Where such evidence exists, comparison between existing substance abuse policies and available interventions shown to be quantifiably effective may indicate the extent to which aggregate costs are avoidable.

As an example, a *Regulatory Impact Analysis Statement* drawn up in support of the proposed Canadian *Cigarette Ignition Propensity Regulation* estimates that a requirement that all cigarettes be self-extinguishing would in Canada reduce fires caused by manufactured cigarettes by between 34 per cent and 68 per cent (*Canada Gazette*, June 29, 2005). This would in turn reduce the number of injuries and deaths, and reduce the amount of property damage by commensurate amounts. This outcome would depend not upon a reduction in exposure to smoking. Thus, this type of policy evaluation may provide information which is not provided by exposure-based data. Similarly, since the result is derived from physical and economic research, rather than from a study of outcomes (since laws requiring cigarettes to be self-extinguishing have not been implemented in any other country), an outcome-based method would also be inappropriate in this case.

There will be circumstances in which exposure-based measures, while relevant, are insufficient on their own to indicate avoidable proportions. Drink-driving is a case in point. While reductions in overall levels of alcohol consumption are likely to lead to a lower prevalence of drink-driving, regulatory enforcement using, for example, low maximum blood alcohol levels, extensive random breath testing and severe penalties has been shown to significantly reduce the prevalence of drink-driving. The most effective interventions often require both reduced exposure and effective enforcement of a regulatory environment.

Reference to evidence on the effectiveness of policy interventions is a very useful addition to the tools of avoidable cost methodology in certain circumstances. Evidence on the effectiveness of a range of interventions is presented in Appendix E.

4 Special considerations in developing countries

Chapter 5 of the aggregate cost *Guidelines* devotes substantial attention to data requirements in developing countries and to special considerations of cost estimation in these countries. The comments in that chapter apply equally to the estimation of avoidable costs, although the development by the WHO of the epidemiological information presented in Ezzati *et al* has ameliorated the data problems of developing countries in a very important respect. One major obstacle to the estimation of the social costs of drug-attributable mortality and morbidity in developing countries has been largely removed.

A major area of cost estimation which still remains largely unresolved relates to the social costs of drug-attributable crime. Although tobacco-related smuggling and tax evasion are major problems in many countries, the major crime related issues which researchers have attempted to address relate to alcohol and, particularly, to illicit drugs. This is an issue which the *Inter-American Drug Abuse Control Commission (CICAD)* has been addressing and Appendix H summarises a presentation which Dr Augusto Pérez-Gómez, Lead Researcher at CICAD, made to the 2005 Ottawa workshop on issues regarding the estimation of the avoidable costs of drug-attributable crime.

5 Time lags associated with policies to reduce substance abuse

Before estimating avoidable proportions it is necessary to specify the time period for which the estimates are to be made. Abuse costs are considered to be avoidable as a result of the application of appropriate anti-abuse policies and programmes. However, some of the identified costs of abuse, while in theory avoidable, may be reduced or eliminated only over long lead times, of which there are three categories:

1. There will be policy implementation lead times, since policies cannot be instantly, designed, legislated and implemented;
2. Even after full and effective implementation there may well be long lead times before the health or other effects of the policies are fully felt;
3. As some abuse costs result from premature mortality, if the prevalence of substance abuse is reduced it will be years before the population structure fully reflects the reduction in premature mortality.

In general, the longer the period of analysis over which the estimates are undertaken, the higher will be the proportion of costs which will be avoidable, although there will certainly be a time period beyond which no further cost reductions are possible. In addition, lead times are likely to differ according to the type of cost under consideration. Taking the example of tobacco, a decline in smoking prevalence may lead to a virtually instant decline in some costs, for example those arising from fire-related deaths, injuries, damage and litter. On the other hand, other types of costs may only be responsive to declines in smoking prevalence with a considerable lag. For instance, reduced smoking prevalence may lead to a decline in lung cancer-related costs only after a period of many years, and the costs of alcohol-attributable road accidents may be borne for many years after the accidents themselves. Acute conditions will generally have shorter lead times, while for chronic conditions the lead times will be much longer.

Researchers will have to make the decision as to whether avoidable proportions should be calculated for a given point in time for a period of time sufficiently long for minimum levels of substance abuse costs to be achieved. It may be possible to produce a time profile of the increase in the proportion of costs which would be avoidable.

In principle, estimates should be made of the lags involved in addressing each type of cost (for example, in relation to health costs, each medical condition). If this is not possible, then an estimate should be made for the average lag time involved in reducing substance abuse. Even this more general approach may, in practice, be impossible to implement. At times, it may be necessary for researchers to make an educated guess as to the range of values in which the actual average lag may lie and to test the sensitivity of the results to the adoption of different lags.

In utilizing avoidable cost estimates in cost-benefit analysis, it will inevitably be necessary to discount the time stream of benefits (that is, reduction in substance abuse costs) to some base year. All other things being equal, the longer the assumed lag the

lower will be the calculated social benefits. Total benefits will be lower and, since they will accrue later in the life of the program, their discounted present value will be less.

The principles involved in consolidating a stream of costs or benefits over time through discounting are discussed in the original *Guidelines*. However, because avoidable costs change over time more than in the standard counterfactual scenario (see Chapter 3), it is necessary here to consider discounting further.

It should be noted that discounting involves adding up the stream of costs or benefits, giving a lower weight to those which occur further in the future. The stream is capitalised into a lump sum called the *net present value* (NPV). Comparisons between the NPV of the benefit and cost streams indicate whether a project is worthwhile. Note, however, that social costs of substance abuse are usually presented as the annual (net) cost, which is assumed to be broadly the same each year. They could be converted to an NPV, but that is not necessary because the streams do not change much through time. This means that an avoidable cost estimate based on an NPV is not comparable to the standard annual cost figure. To make them comparable, the NPV should be multiplied by the discount rate to convert the capitalised sum into the (weighted) average cost or benefit stream that the NPV generates.

6 The reliability and usefulness of avoidable cost estimates

The estimation of the avoidable costs of substance abuse is a two stage process. The first stage is to estimate the total social costs of substance abuse. The second is to estimate the proportions of each category of the total costs which are potentially avoidable. Potential errors arise at each of these two stages.

6.1 Problems with estimating total costs

The following discussion is drawn from Single and Easton (2001).

Although the development of international guidelines has done much to increase comparability of results, there remains a lack of consensus regarding the appropriate methodology to employ in conducting cost estimation studies. The more commonly used Cost-of-Illness (COI) approach has been criticized for including indirect costs such as productivity costs, and some prefer the more conservative "externality" approach championed by US economist Willard Manning. Even among those who employ a COI approach, there are differences regarding the valuation of premature mortality caused by substance abuse. While most studies continue to use the human capital approach (which uses forgone income to estimate forgone productivity), Collins and Lapsley have used the alternative "demographic" approach and several studies use new "willingness-to-pay" techniques. For example, tobacco costs have been estimated using this latter technique by Brian Easton in New Zealand (Easton, 1997) and Claude Jeanrenaud in Switzerland (Jeanrenaud *et al*, 1999).

In almost all cost studies some cost data are lacking. This is particularly true for developing economies which often lack reliable reporting systems. Even in developed economies, there is often sparse information on many cost elements. For example, estimates of the proportions of the various types of crime attributable to substance abuse can be highly contentious. There tends to be a lack of data on the costs of specific drug-related productivity problems such as absenteeism, job turnover, lower on-the-job productivity, drug-attributable disability and so forth. In some countries, estimates of substance use do not exist for the year under investigation, and must be calculated by interpolating the prevalence from other years. It is frequently difficult to determine from government budgets what proportion of policy costs (prevention, research and law enforcement costs) that can be attributed to particular drugs of abuse. In the absence of complete data, judgments have to be made, because to ignore a cost is in effect to count that cost as zero. To ignore a cost completely is generally more erroneous than making a judgment on the basis of incomplete information

Even when relatively complete data are available, the prevailing methods for estimating economic costs attributable to substance abuse involve a layering of multiple assumptions. For example, estimates of alcohol-attributable morbidity and mortality are required to underpin estimates of productivity costs and costs to the health care system. Morbidity and mortality estimates are made by combining information on (a) the relative risk of consuming alcohol at different levels to various causes of disease and death from

meta-analyses of the epidemiological literature with (b) prevalence data on the number of persons consuming alcohol at levels associated with a higher relative risk in order to generate (c) aetiologic fractions of the proportion of all such causes of disease and death that can be causally ascribed to alcohol use. These aetiologic fractions are then applied to the reported number of hospitalizations and deaths categorized by cause to estimate the morbidity and mortality attributable to alcohol use. This procedure must necessarily make the following assumptions:

- All alcohol-related causes are included in the epidemiological data which form the basis for the cost calculations.
- Where some conditions are related to both alcohol and other causes (for example, injuries stemming from fires involving both smoking and alcohol intoxication), a correct division of attribution can be made.
- Relative risk estimates from studies in one country can be used to estimate relative risk in another in which there are no local data.
- Confounders are adequately controlled for in the studies used to estimate relative risk.
- Age and gender are adequately controlled for in estimating relative risk.
- Since, in most cases, the estimates of relative risk are derived from both morbidity and mortality studies, the risk of morbidity is equivalent to the risk of mortality.
- The reported number of hospitalizations and deaths is accurately counted and complete, and the causes are accurately recorded. It is known, for example, that some disorders only recently described in the medical literature (such as fetal alcohol syndrome) are not yet being reliably recorded.

Another major reason for caution in interpreting the bottom-line cost estimates concerns changes in the epidemiological database and what is known about the effects of substance abuse. Improvements in diagnostic practices are constantly occurring. For example, the best method currently available for estimating alcohol-attributable morbidity and mortality relies on reliable diagnoses of alcohol-related causes of death and hospitalization. Conditions such as fetal alcohol syndrome have only recently been described and accepted in the medical literature, and such conditions will likely be underreported for some time. More importantly, new research is continually emerging concerning the link between substance abuse and various causes of disease and death. Ten years ago, there were insufficient data to conclude that there is a causal connection between alcohol use and breast cancer. Now the evidence is compelling. Although alcohol accounts for less than 3% of breast cancer fatalities in Canada, the numbers are so large that it represents the third leading cause of alcohol-attributable death among women. A study conducted just one decade earlier would likely have not even included breast cancer in the cost calculations.

6.2 Difficulties in estimating avoidable proportions

Avoidable costs represent a proportion of the total social costs of substance abuse. The process of estimating the value of avoidable costs involves estimating the avoidable proportions of all the individual categories of substance abuse costs and then applying these proportions to the aggregate cost estimates. Thus, all of the above problems and sources of error apply to the estimation of avoidable costs, as well as to the estimation of total costs.

In addition, specific problems arise in estimating the avoidable fractions. If we had access to perfect information, we would need, in addition to robust estimates of all cost elements from a cost estimation study:

- detailed data on risk factors, for example the relative contributions of different levels of drinking, smoking and use of illicit drugs to each problem indicator;
- detailed information on the time lags between engaging in the risky behaviour and the onset of the attributable problem.

Estimation of the feasible minimum for each category of cost raises its own problems. Choosing a feasible minimum may well involve choice of a reference country. It might be possible simply to choose a reference country for each major category of cost – health care, productivity, crime, pollution etc., but in practice there will probably be a whole range of cost sub-categories for which the choice of reference countries will be necessary. Best practice performance for the cost category of one drug may not be achieved, or achievable, in the same country as for another cost category of the same drug. This is likely to be particularly true in relation to illicit drugs, since this term covers a range of diverse drugs with different consumption prevalence, cost impacts and reaction to policy initiatives (for example, to policies to reduce demand or supply).

As mentioned earlier, the unavoidable cost proportions will depend upon the time periods chosen for the analysis, declining over time until some irreducible minimum is reached. Determination of a feasible minimum from international comparisons of outcomes will indicate what this irreducible minimum is likely to be but not the time period taken to achieve it. Given the paucity of information on these time lags, it may be necessary to assume, for all cost categories for a particular drug, a common lag long enough to permit reasonable certainty that all potentially avoidable costs can, in fact, be avoided.

6.3 The benefits of estimating avoidable costs

Even with improvements in methods for estimating drug-related mortality, morbidity and economic costs, and their avoidable components, there are still significant sources of error in the bottom-line estimates of total economic costs caused by substance abuse. Despite these uncertainties, it is considered that it is, for the following reasons, a useful exercise to undertake studies of the avoidable cost of substance abuse.

The first reason is that undertaking the exercise exposes data deficiencies, and forces us to improve the statistical base and our understanding of the processes involved. Economic cost studies help to identify information gaps, research needs and desirable refinements to

national statistical reporting systems. There is no better way to lay out a national research agenda than to conduct a cost estimation study. Cost studies are an excellent device for the identification of data development and research needs. For example, Canadian and Australian cost studies identified a strong need for improved estimation of the proportion of crime that can be attributed to alcohol and drug misuse and spawned studies in both these countries.

The second response concerns quality control. Policy makers need and use cost estimates, explicitly or implicitly, in setting priorities among competing concerns. Despite all of the uncertainties involved, economic cost estimates are frequently used to argue that policies on alcohol and other psychoactive substances should be given a high priority on the public policy agenda. The public is entitled to a quality standard against which individual cost estimation studies can be assessed. Without such a standard there will be a tendency by the advocates for each social problem to overbid, adding in additional items to make their concern a suitably high (even exaggerated) number.

Estimates of the aggregate costs of substance abuse do not, in themselves, indicate the value of the potential returns to drug intervention strategies. The potential returns, which are measured as the *avoidable* costs of substance abuse, represent more meaningful estimates of potential public policy benefits of interventions. They can, therefore, represent a very useful tool in putting the case for the allocation or greater public resources to drug programs. As an example, Appendix G presents the steps which researchers might have to take when attempting to estimate the value of the social benefits (that is, the reduction in social costs) which would result from a specified reduction in substance abuse, for example a specified reduction in smoking prevalence.

Thirdly, cost estimates help to appropriately target specific problems and policies. It is important to know which aspects of substance abuse involve the greatest economic costs, what specific problems are most likely to occur and in what demographic or geographic groups. The nature and magnitude of costs draw our attention to specific areas which need public attention, or where specific measures may be effective.

Further, the estimation of avoidable costs involves the gathering of information on international experience in the effectiveness of a wide range of drug interventions. As part of the process of estimating avoidable costs, a considerable amount of such information is likely to be gathered and collated, assisting in the design of improved and efficient drug programs. Determination of the avoidable costs proportion does not, of itself, indicate the policy mix by which the irreducible minimum of costs may be achieved. However, inspection of the policy mixes in the jurisdictions representing feasible minimums may give a strong indication of appropriate policies.

Last but not least, the development of improved estimates of the avoidable costs of substance abuse offers the potential, although generally not yet realised, to provide baseline measures for more sophisticated economic analyses to determine which policies and programmes are the most effective in reducing the harm associated with alcohol and other drug use. A combined total cost/avoidable cost study would provide a high proportion of the data requirements for detailed program analysis.

The concept of Gross Domestic Product (GDP) was subject to much the same sources of error and criticisms that aggregate economic cost estimates currently face, and that avoidable cost estimates are likely to face. Despite similar issues of lack of complete data and layering of assumptions, the estimation of GDP has been continually refined and improved, and it has become one of the most useful tools for economic analysis and policy development. Through international cooperation and the development of an on-going process to continually update and refine the methodology, avoidable cost studies can similarly become more reliable over time, and perhaps even become an important cornerstone for comparative analyses of substance abuse policy and interventions much as GDP is today.

7 Policy implications of avoidable cost estimates

7.1 Complexities of policies to reduce substance abuse

In comparison with alcohol and illicit drugs, the design of policy to minimise tobacco-attributable harm is relatively straightforward, for three reasons:

- Since any non-trivial consumption of tobacco is harmful to health, the objective of public policy should be the simple one of reducing the use of tobacco by the maximum amount possible.
- Tobacco is a relatively homogeneous product. Admittedly, as well as cigarettes/cigars, it can take the form of oral or chewing tobacco. However these latter two uses represent, in most countries, a small proportion of total tobacco consumption and their chronic impacts are similar to those of smoking.
- Since tobacco is usually a legal product (though not necessarily when used by minors), excise taxes and advertising bans (both powerful tools) can be used to reduce consumption.

Illicit drugs represent a more complex policy issue, for two main reasons:

- A whole range of products with diverse health and other effects is subsumed within the title of illicit drugs. Policy appropriateness can vary from drug to drug.
- As a result of the illegality, by definition, of these drugs, a range of polices, including taxation and regulation of product quality, are precluded, and the underlying data are much less robust.

Alcohol, too, is not a homogeneous product, with the three major product categories of beer, wine and spirits. A further consideration is that, under certain conditions, the consumption of alcohol has been proven to have protective effects.

7.2 Dealing with the protective effects of alcohol

In calculating both the aggregate and the avoidable costs of alcohol use, a complication arises which does not exist in the cases of tobacco or illicit drugs. There is significant evidence that, for some medical conditions, alcohol consumption at appropriate levels can have a protective effect, that is, alcohol consumption can reduce the risk of illness or death. In these circumstances, abstinence-based attributable fractions (as opposed to attributable fractions based on hazardous or harmful consumption) are negative. With very minor exceptions, there is no evidence of any analogous health benefits from consumption of tobacco or illicit drugs.

This is an issue that has been addressed by, *inter alia*, Ridolfo and Stevenson (2001) and Rehm *et al* (2004). They conclude that low-to-moderate levels of alcohol consumption can, given appropriate drinking patterns, confer health benefits in relation to ischaemic heart disease, supraventricular cardiac dysrhythmias, cholelithiasis (gallstones), ischaemic strokes, haemorrhagic strokes, hypertension and type II diabetes.

The existence of protective effects of alcohol raises the issue of whether these protective benefits (which can be considered as negative costs) should be incorporated in estimates of the aggregate social costs of alcohol, and therefore also in avoidable cost estimates. Some aggregate cost studies (for example Collins and Lapsley, 2002) have taken these benefits into account, while for other researchers (for example Easton, 1997) the inclusion of the benefits of alcohol "misuse" was considered inappropriate.

This issue is complicated by the fact that it appears that even so-called "responsible' levels of alcohol consumption can be dangerous in relation to certain medical conditions, for example female breast cancer. Furthermore, even risky/high risk drinking can prevent some deaths (while causing far more). There is, thus, a difficulty in defining the term alcohol "misuse". How can alcohol consumption which saves some lives (while also causing deaths) be considered to be misuse in all cases. Conversely, how can some level of alcohol consumption be considered to be "safe" when it causes some deaths (while preventing others)?

This point is illustrated in a paper by Chikritzhs, Stockwell *et al* (2002) who estimate the numbers of lives lost and lives saved in Australia in 1998 due to low risk and risky/high risk drinking, compared with a baseline of complete abstinence. Their results are presented in Table 9 below.

Table 9 – Estimated numbers of lives lost and saved due to low risk and risky/high risk drinking when compared to abstinence, Australia, 1998

	Low risk drinking	Risky/high risk drinking	All drinking
Lives lost	1,505	3,294	4,799
Lives saved	(6,605)	(557)	(7,162)
Total	(5,100)	2,737	(2,363)

Note: figures in brackets represent numbers of lives saved.

Source: Chikritzhs, Stockwell *et al* (2002), Table 1

They conclude that:

It is recommended that for future reports on alcohol-caused morbidity and mortality, there would be value in presenting a more detailed picture that identifies both the costs and the benefits of low risk drinking and risky/high risk drinking. In order to do this, an abstinence-based contrast must be adopted.

If protective benefits are taken into account in aggregate cost estimates they should also be incorporated in avoidable cost estimates. However, policies designed to minimise the costs of alcohol use may also reduce its benefits. The existence of protective benefits throws doubt on the usefulness of the prevention paradox (Kreitman, 1986) as a guide to alcohol policy. Policies aimed at minimising the costs of alcohol misuse (for example, high excise taxes on alcohol) may, in the process, reduce the number of low or moderate

alcohol consumers, and so reduce the total benefits of low or moderate alcohol consumption.

Thus, measures of the Arcadian normal appropriate for estimating avoidable costs may differ completely from those appropriate for alcohol benefits. In practical policy terms, the best outcome to be hoped for may be to reduce alcohol costs while maintaining alcohol benefits unaffected.

7.3 Policies available to reduce substance abuse costs

A very broad range of measures is available to minimise the social costs of substance abuse. Loxley *et al* (2004), in surveying the effectiveness of potential prevention policies, categorized available measures under the following broad headings:

- Interventions for children (0-11 years);
- Interventions for young people (12-24 years);
- Broad-based prevention;
- Demand reduction;
- Regulation and law enforcement: licit drugs;
- Regulation and law enforcement: illicit drugs;
- Judicial procedures; and
- Harm reduction interventions.

Appendix E provides greater detail of the available interventions and of the authors' judgment of the effectiveness of these interventions. Further information about the effectiveness of interventions to reduce drug-attributable crime is presented in Appendix D.

Reviews such as that provided by Loxley *et al* (2004) provide an important basis for the development of effective intervention strategies.

8 Conclusion

This study addresses the provision of guidelines for estimation in an area which represents an almost completely new area of research. We should underestimate neither the importance of such estimation, nor their difficulties, nor the interpretation of the results. In principle, the aim should be a "gold standard" methodology which is solidly grounded in epidemiological and economic theory, and one for which sound data are available. Such gold standard research would yield results which could be applied to maximise the quality of public policies designed to deal with the problem of substance abuse.

In practice, it is never easy to apply a gold standard. This was clearly acknowledged in the development of the original aggregate cost *Guidelines* and it is certainly the case in the estimation of avoidable costs. In the latter case it is necessary both to develop the necessary theory and to review and improve available data. It is also important to clarify the theoretical distinction between *minimum achievable* levels of substance abuse and *optimal* levels of abuse. These guidelines are concerned with the estimation of feasible minimum levels, not optimal levels.

This study reviews three broad possible approaches to the estimation of avoidable costs. The first method involves what might be called a classic epidemiological approach. This is probably the most theoretically correct of the three, being exposure-based, but also is the one which has the most challenging data requirements. A second approach is based on what has become known as the Arcadian normal, which uses best outcomes (in terms of morbidity and/or mortality in comparable countries) as the basis for estimating avoidable proportions. It was, however, agreed at the Ottawa workshop that an exposure-based methodology was superior to one based on outcomes. The desirability of an exposure-based approach applied in a situation of data deficiency led to a third methodology using exposure-based comparators, drawing on international data recently published by the WHO on attributable fractions.

All three methodological approaches addressed here relate to using morbidity and mortality cost calculations to estimate the Feasible Minimum. Studies of the avoidable costs of crime are, in future, likely to reveal further complexities which will require careful analysis and interpretation.

There are also likely to be benefits from the examination of international evidence on the effectiveness of interventions designed to deal with the levels, or adverse effects, of substance abuse. Where strong evidence exists about the effectiveness of particular interventions, comparison of actual policy mixes with interventions known to be effective in other countries may help to indicate the proportions of substance abuse costs which are avoidable.

Research on avoidable costs may need to consist of a hybrid of more than one, and possibly all, of the above approaches. It should be acknowledged that the development of the methodology of estimation of avoidable costs is work-in-progress. It is expected that the publication of these guidelines will encourage pilot studies in a range of countries and

for a range of drugs. The results of such studies should inform the research experience, which in turn will permit the improvement and refinement of these guidelines.

Bottom-line avoidable cost estimates are not the sole consideration in the determination of political priorities, nor should they be. There are a host of other considerations that policy makers and others involved in policy development should consider. However, economic cost studies provide an ideal framework for identifying key leverage points in policy development and for the development of priorities in research as well as for treatment, prevention and other interventions. The cost estimation symposia and international guidelines have done much to reduce differences in economic modelling and enhance the comparability of results of cost estimation studies in different countries. The current process of developing and refining avoidable cost guidelines will, hopefully, have a similar positive result. Despite the many sources of error in current cost estimates, it can be argued that it is indeed a valuable research exercise to estimate the costs of substance abuse and continually improve upon the methodology for conducting such studies. And ultimately, if appropriately qualified researchers do not do these cost studies well, others will do them badly, compounding the sorts of problems described here, while failing to provide the sort of benefits that good studies can do, and so simply adding to the confusion in a field which is already very difficult.

References

1. American Psychiatric Association (1994), Diagnostic and Statistical Manual of Mental Disorders. Fourth ed., Washington D.C.

2. Andlin-Sobocki P. and Rehm J. (2005), "Cost of addiction in Europe", *European Journal of Neurology*, 12(s1): 28-33.

3. Australian Bureau of Statistics, *Corrective Services, Australia* (4512.0).

4. Australian Bureau of Statistics, *Crime and Safety Australia*, (4509.0).

5. Australian Bureau of Statistics, *Prisoners in Australia* (4517.0).

6. Australian Bureau of Statistics, *Recorded Crime*, *Australia* (4510.0).

7. Armstrong, B.K. (1990), "Morbidity and mortality in Australia: How much is preventable?", in John McNeill *et al* (eds.), *A Handbook of Preventative Medicine,* Edward Arnold.

8. Babor, T.; Caetano, R,; Casswell, S.; Edwards, G.; Giesbrecht, N.; Graham, K.; Grube, J.; Gruenewald, P.; Hill, L.; Holder, H.; Homel, R.; Österberg, E.; Rehm, J.; Room, R.; and Rossow, I. (2003), *Alcohol: no ordinary commodity. Research and public policy.* Oxford and London: Oxford University Press.

9. Barengo, N.C.; Sandstrom, P.H.; Jormanainen, V.J.; and Myllykangas, M.T. (2004), "Changes in smoking prevalence among Finnish physicians 1990-2001", *European Journal of Public Health*, June; 14(2):201-3.

10. Bryant, Meredith and Williams, Paul (2000), "Alcohol and Other Drug-Related Violence and Non-Reporting", *Trends and issues in crime and criminal justice*, No. 171, October, Australian Institute of Criminology.

11. Carcach, Carlos (1997), "Reporting Crime to the Police", *Trends and issues in crime and criminal justice*, No. 68, March, Australian Institute of Criminology.

12. Carcach, Carlos and Grant, Anna (2000), "Imprisonment in Australia: The Offence Composition of Australian Correctional Populations, 1988 and 1998", *Trends and issues in crime and criminal justice*, No. 164, July, Australian Institute of Criminology.

13. Carcach, Carlos and McDonald, David (1997), *National Police Custody Survey August 1995*, Australian Institute of Criminology Research and Public Policy Series No. 9.

14. Chikritzhs, Tanya; Stockwell, Tim; Jonas, Helen; Stevenson, Chris; Cooper-Stanbury, Mark; Donath, Susan; Single, Eric; and Catalano, Paul (2002), "Towards a standardised methodology for estimating alcohol-caused death, injury and illness in Australia", *Australian and New Zealand Journal of Public Health* 26 (5).

15. Coddle, J. (2002), "Appendix B, Method of estimation of drug-attributable cases", in Collins and Lapsley (2002).

16. Collins, D.J. and Lapsley, H.M.(1999), *The Social Costs of Tobacco in Victoria and the Social Benefits of Quit Victoria*, Victorian Smoking and Health Program, Melbourne.

17. Collins, D.J. and Lapsley, H.M. (1996), *The social costs of drug abuse in Australia in 1988 and 1992*, National Drug Strategy Monograph Series No. 30, http://www.health.gov.au/pubhlth/publicat/document/mono30.pdf.

18. Collins, D.J. and Lapsley, H.M. (2002), *Counting the cost: estimates of the social costs of drug abuse in Australia in 1998-9*, National Drug Strategy Monograph Series No. 49, http://www.health.gov.au/pubhlth/publicat/document/mono49.pdf.

19. Collins, D.J. and Lapsley, H.M (2004), *Counting the costs of tobacco and the benefits of reducing smoking prevalence in Western Australia,* The Cancer Council Western Australia, Monograph Series No. 4 2004.

20. Degenhardt, Louisa; Hall, Wayne; Warner-Smith, Matthew; and Lynskey, Michael (2004), "Illicit drug use", in Ezzati *et al* (2004).

21. Easton, Brian (1997), "The Social Costs of Tobacco Use and Alcohol Misuse", report prepared for the Alcohol Advisory Council of New Zealand, the Health Research Council of New Zealand and the Public Health Commission, *Public Health Monograph No. 2*, Department of Public Health, Wellington School of Medicine.

22. Eide, G.E. and Heuch, I. (2001), "Attributable Fractions: Fundamental Concepts and Their Visualization", *Statistical Methods in Medical Research*,10: 159–193.

23. Ezzati, Majid; Lopez, Alan D.; Rodgers, Anthony; and Murray, Christopher J.L. (Editors) (2004), *Comparative Quantification of Health Risks, Volume 1*, World Health Organization.

24. Ezzati, Majid and Lopez, Alan D. (2004), "Smoking and oral tobacco use", in Ezzati *et al* (2004).

25. Gmel, G.; Gutjahr, E.; and Rehm, J. (2003), "How stable is the risk curve between alcohol and all-cause mortality and what factors influence the shape? A precision-weighted hierarchical meta-analysis", *European Journal of Epidemiology*, 18(7): 631-642.

26. Goldstein, P.J. (1985), "The drugs/violence nexus: A tripartite conceptual framework", *Journal of Drug Issues*, 14, 493-506.

27. Greenland, S. and Robins, J.M. (1988), "Conceptual Problems in the Definition and Interpretation of Attributable Fractions", *American Journal of Epidemiology*, 128: 1185–1197.

28. Harwood, H.; Fountain, D.; and Livermore G, (1998), *The Economic Costs of Alcohol and Drug Abuse in the United States 1992*, Report prepared for the National Institute

on Drug Abuse and the National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Department of Health and Human Services, NIH Publication No. 98-4327. Rockville, MD: National Institutes of Health. http://www.nida.nih.gov/EconomicCosts/Index.html

29. Jeanrenaud, C., Priez, F. and Vitale, S. (1999), "The social cost of smoking in Switzerland", in Jeanrenaud and Soguel (eds.) (1999).

30. Jeanrenaud, C. and Soguel, N. (1999), Valuing the Cost of Smoking: Assessment Methods, Risk Perception and Policy Options, Kluwer Academic Publisher, Studies in Risk and Uncertainties, Boston. Kreitman, N. (1986), Alcohol consumption and the prevention paradox", British Journal of Addiction, 81.

31. Kreitman, N. (1986), *Alcohol consumption and the prevention paradox*", British Journal of Addiction, 81.

32. Loxley, Wendy; Toumbourou, John W.; Stockwell, Tim; Haines, Ben; Scott, Katie; Godfrey, Celia; Waters, Elizabeth; Patton, George; Fordham, Richard; Gray, Dennis; Marshall, Jann; Ryder, David; Saggers, Sherry; Sanci, Lena; and Williams, Jo (2004a), *The Prevention of Substance Use, Risk and Harm in Australia: a review of the evidence*, Ministerial Council on Drug Strategy Monograph.

33. Loxley, Wendy; Toumbourou, John W.; and Stockwell, Tim (2004b), *Summary: The Prevention of Substance Use, Risk and Harm in Australia: a review of the evidence*, Ministerial Council on Drug Strategy.

34. MacDonald, Ziggy; Tinsley, Louise; Collingwood, James; Jamieson, Pip and Pudney, Stephen (2005), *Measuring the harm from illegal drugs using the Drug Harm Index*, Home Office Online Report 24/05, available at http://www.homeoffice.gov.uk/rds/pdfs05/rdsolr2405.pdf

35. Mackay, J. and Eriksen, M. (2002), *The Tobacco Atlas*, World Health Organisation.

36. Makkai, Toni (1999c), *Drug Use Monitoring in Australia (DUMA): a brief description*, Australian Institute of Criminology, Research and Public Policy Series, no 21.

37. Makkai, Toni (2000), "Patterns of Drug Use Amongst Police Detainees: 1999-2000", *Trends and Issues in Crime and Criminal Justice*, No 185, December , Australian Institute of Criminology.

38. Makkai, Toni and McGregor, Kiah (2001), *Drug Use Monitoring In Australia: 2000 Annual Report on Drug Use Among Police Detainees*, Australian Institute of Criminology Research and Public Policy Series, No. 37.

39. Makkai, Toni and McGregor, Kiah (2002), *Drug Use Monitoring In Australia: 2001 Annual Report on Drug Use Among Police Detainees*, Australian Institute of Criminology Research and Public Policy Series, No. 41.

40. Maldonado, G. and Greenland, S. (2002), "Estimating Causal Effects", *International Journal of Epidemiology,* 31: 422–429.

41. Mathers, Colin D.; Bernard, Christina;, Iburg, Kim Moesgaard; Inoue, Mie; Fat, Doris Ma; Shibuya, Kenji; Stein, Claudia; Tomijima, Niels; and Xu, Hongyi (2003), *Global Burden of Disease in 2002: data sources, methods and results*, Global Programme on Evidence for Health Policy Discussion Paper No.54, World Health Organization, December.

42. Mathers CD, Stein C, Ma Fat D et al. (2002) *The global burden of disease 2000: version 2, methods and results.* (Discussion Paper No. 50.) Evidence and Information for Policy, World Health Organization, Geneva. Also available at http://www.who.int/evidence.

43. McFadden, Michael and Mwesigye, Sue-Ellen (2004a), "AFP Drug Harm Index", Australian Federal Police conference paper

44. McFadden, Michael and Mwesigye, Sue-Ellen (2004b), "Drug Harm Index-Revised", *Platypus Magazine*, October, also available at www.afp.gov.au/afp/raw/publications/platypus/oct04/45-48.pdf

45. Miettinen, O.S. (1972), "Standardization of risk ratios", *American Journal of Epidemiology*, 96(6): 383-388.

46. Murray, C.J. and Lopez, A.D. (1997), "Mortality by cause for the eight regions of the world: the global burden of disease study", *Lancet*, 349(9061): 1269-1276.

47. Murray, C.J. and Lopez, A.D. (1999), "On the comparable quantification of health risks: lessons from the Global Burden of Disease Study",. *Epidemiology*, 10(5): 594-605.

48. Murray, C.J.; Ezzati, M.; Lopez, A.D.; Rodgers, A.; and VanderHoorn, S (2003). "Comparative quantification of health risks conceptual framework and methodological issues", *Population Health Metrics*, 1(1): 1-20.

49. Perez, A.; Wilson, E.; and Valencia, J.E. (2004), *Program to estimate the human, social, and economic costs of drugs in the Americas – Research Manual*, Inter-American Observatory on Drugs, Inter-American Drug Abuse Control Commission (CICAD).

50. Pernanen, K.; Brochu, S.; Cousineau, M.-M.; Cournoyer, L.G.; and Fu Sun (2000), "Attributable fractions for alcohol and illicit drugs in relation to crime in Canada: conceptualization, methods and internal consistency of estimates", *Bulletin on Narcotics*, Volume LII, Nos. 1 and 2, United Nations Drug Control Programme.

51. Pernanen, K.; Cousineau, M-M; Brochu, S.; and Fu Sun (2002), "Proportions of Crimes Associated with Alcohol and Other Drugs in Canada", Canadian Centre on Substance Abuse, www.ccsa.ca/docs/crime2002.pdf.

52. Rehm, J.; Gutjahr, E.; and Gmel, G. (2001), "Alcohol and all-cause mortality: a pooled analysis", *Contemporary Drug Problems*, 28: 337-361.

53. Rehm, J. (2003), "Suchtmittel und Public Health", *Suchttherapie*, 4: 72-75.

54. Rehm, J.; Room, R.; Graham, K.; Monteiro, M.; Gmel, G.; and Sempos, C.T. (2003a). "The relationship of average volume of alcohol consumption and patterns of drinking to burden of disease - an overview", *Addiction*, 98(10): 1209-1228.

55. Rehm, J.; Sempos, C.T.; and Trevisan, M. (2003b), "Average volume of alcohol consumption, patterns of drinking and risk of coronary heart disease – a review", *Journal of Cardiac Risk*, 10 (1): 15-20.

56. Rehm, Jürgen; Room, Robin; Monteiro, Maristela; Gmel, Gerhard; Graham, Kathryn; Rehn, Nina; Sempos, Christopher, T.; Frick, Ulrich; and Jernigan, David (2004), "Alcohol use", in Ezzati *et al* (2004).

57. Ridolfo, Bruno and Stevenson, Chris (2001), *The quantification of drug-caused mortality and morbidity in Australia, 1998*, Australian Institute of Health and Welfare, Drug Statistics Series No. 7.

58. Room, R.; Babor, T.; and Rehm, J. (2005), "Alcohol and public health", *Lancet*, 365(9458): 519-30.

59. Rothman, K.J. and Greenland, S. (1998), *Modern Epidemiology*, Philadelphia: Lippincott-Williams.

60. Single, E.; Collins, D.; Easton, B.; Harwood, H.; Lapsley, H.; and Maynard, A. (1996), *International Guidelines for Estimating the Costs of Substance Abuse*, Canadian Centre on Substance Abuse

61. Single, E. and Easton, B. (2001), *Estimating the economic costs of alcohol misuse: why we should do it even though we shouldn't pay too much attention to the bottom line result*s, paper presented at the annual meeting of the Kettil Bruun society for Social and Epidemiological research on Alcohol, Toronto, May.

62. Single, E.; Collins, D.; Easton, B.; Harwood, H.; Lapsley, H.; Kopp, P.; and Wilson, E (2003), *International Guidelines for Estimating the Costs of Substance Abuse-Second Edition*, World Health Organisation.

63. Skog, O.J. (1985), "The collectivity of drinking cultures: a theory of the distribution of alcohol consumption", *British Journal of Addiction*, 80(1): 83-99.

64. Stevenson, Richard J. and Forsythe, Lubica M.V. (1998), *The Stolen Goods Market in New South Wales. An interview study with imprisoned burglars*, New South Wales Bureau of Crime Statistics and Research.

65. Taylor, B.; Rehm, J.; and Gmel, G. (2005), "Moderate Alcohol Consumption and the Gastrointestinal Tract", *Journal of Dig Dis*, in press.

66. Walter, S.D. (1976), The estimation and interpretation of attributable risk in health research", *Biometrics* , 32: 829-849.

67. Walter, S.D. (1980), "Prevention of multifactorial disease", *American Journal of Epidemiology*, 112: 409–416.

68. Warner, Kenneth E. and Burns, David M. (2003), „Hardening and the hard-core smoker: concepts, evidence, and implications", *Nicotine and Tobacco Research*, 5, 37-48.

69. World Health Organization (2002), *The World Health Report 2002: Reducing Risks, Promoting Healthy Life*, Geneva.

70. Young, J.M. and Ward, J.E. (1997), "Declining rates of smoking among medical practitioners", *Medical Journal of Australia*, August 18; 167(4):232.

Appendix A – World Health Organisation information on drug-attributable fractions

As indicated earlier, since the publication of the original aggregate cost *Guidelines* the available epidemiological information on the impact of substance abuse worldwide has improved very substantially. This has resulted from the publication by the World Health Organisation of *Comparative Quantification of Health Risks* (Ezzati *et al* 2004). The estimation of the aggregate costs of substance abuse has, as a result, become very much easier for countries which are unable to generate their own epidemiological information.

Table 11, which appears later in this Appendix, provides an example of the comprehensive epidemiological information provided in the Annex to Ezzati *et al* (2004). This information is available on CD-ROM as both PDF and Excel files. The information is provided individually for 14 sub-regions, which are defined in Table 10.

Table 10 – The 14 epidemiological sub-regions

WHO region	*Mortality stratum*[a]	*Countries*
AFR	D	Algeria, Angola, Benin, Burkina Faso, Cameroon, Cape Verde, Chad, Comoros, Equatorial Guinea, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Madagascar, Mali, Mauritania, Mauritius, Niger, Nigeria, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Togo
	E	Botswana, Burundi, Central African Republic, Congo, Côte d'Ivoire, Democratic Republic of the Congo, Eritrea, Ethiopia, Kenya, Lesotho, Malawi, Mozambique, Namibia, Rwanda, South Africa, Swaziland, Uganda, United Republic of Tanzania, Zambia, Zimbabwe
AMR	A	Canada, Cuba, United States of America
	B	Antigua and Barbuda, Argentina, Bahamas, Barbados, Belize, Brazil, Chile, Colombia, Costa Rica, Dominica, Dominican Republic, El Salvador, Grenada, Guyana, Honduras, Jamaica, Mexico, Panama, Paraguay, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Suriname, Trinidad and Tobago, Uruguay, Venezuela
	D	Bolivia, Ecuador, Guatemala, Haiti, Nicaragua, Peru
EMR	B	Bahrain, Cyprus, Iran (Islamic Republic of), Jordan, Kuwait, Lebanon, Libyan Arab Jamahiriya, Oman, Qatar, Saudi Arabia, Syrian Arab Republic, Tunisia, United Arab Emirates
	D	Afghanistan, Djibouti, Egypt, Iraq, Morocco, Pakistan, Somalia, Sudan, Yemen
EUR	A	Andorra, Austria, Belgium, Croatia, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Luxembourg, Malta, Monaco, Netherlands, Norway, Portugal, San Marino, Slovenia, Spain, Sweden, Switzerland, United Kingdom
	B	Albania, Armenia, Azerbaijan, Bosnia and Herzegovina, Bulgaria, Georgia, Kyrgyzstan, Poland, Romania, Serbia and Montenegro, Slovakia, Tajikistan, The former Yugoslav Republic of Macedonia, Turkey, Turkmenistan, Uzbekistan
	C	Belarus, Estonia, Hungary, Kazakhstan, Latvia, Lithuania, Republic of Moldova, Russian Federation, Ukraine
SEAR	B	Indonesia, Sri Lanka, Thailand
	D	Bangladesh, Bhutan, Democratic People's Republic of Korea, India, Maldives, Myanmar, Nepal
WPR	A	Australia, Brunei Darussalam, Japan, New Zealand, Singapore
	B	Cambodia, China, Cook Islands, Fiji, Kiribati, Lao People's Democratic Republic, Malaysia, Marshall Islands, Micronesia (Federated States of), Mongolia, Nauru, Niue, Palau, Papua New Guinea, Philippines, Republic of Korea, Samoa, Solomon Islands, Tonga, Tuvalu, Vanuatu, Viet Nam

[a] **A**: very low child mortality and very low adult mortality; **B**: low child mortality and low adult mortality; **C**: low child mortality and high adult mortality; **D**: high child mortality and high adult mortality; **E**: high child mortality and very high adult mortality. High-mortality developing sub-regions: AFR-D, AFR-E, AMR-D, EMR-D and SEAR-D. Low-mortality developing sub-regions: AMR-B, EMR-B, SEAR-B, WPR-B. Developed sub-regions: AMR-A, EUR-A, EUR-B, EUR-C and WPR-A. This classification has no official status and is for analytical purposes only.

The definitions of what are "very low", "low", "high" and "very high", while arguably somewhat arbitrary and context dependent, are meant to distinguish between countries where the epidemiological transition has essentially been completed (e.g. western Europe, Japan, USA) and those where it has not (e.g. Latin America, China), specifically identifying countries where major health reversals affecting adults have occurred, either because of marked increases in vascular and respiratory diseases and injuries (eastern Europe), or because of HIV/AIDS (southern and eastern Africa).

Source of table and quotation: "Introduction to the Annex Tables", in Ezzati *et al* (2004), Volume 3 (on CD-ROM).

The first three tables appearing in the Annex to Ezzati *et al* (2004) for each risk factor–disease pair present details of the population attributable fractions (PAF) for mortality, years of life lost (YLL) and disability-adjusted life years (DALY) due to each risk factor in 2000, by age, sex and sub-region. The population attributable fraction is the proportional reduction in population disease or mortality that would occur if exposure to the risk factor were reduced to an alternative (counterfactual) exposure scenario. For each risk factor–outcome pair, PAF (ranging from 0 to 100—presented as percentage points) is the best estimate of the *full* effects of exposure on the specific health outcome (disease or injury), either directly or indirectly where the exposure also acts through other, more proximal exposures in a causal chain leading to disease or injury.

The second set of three tables per risk factor–disease pair give the total number of deaths, YLLs and DALYs due to each risk factor in 2000, by age, sex and sub-region. These estimates are obtained by applying the respective population attributable fractions in the first three tables to the estimated total number of deaths, YLL and DALYs for each regional population, estimated as part of the broader GBD 2000 project (Mathers *et al*, 2002). As such, the second set of three tables show the total (absolute) population health effects of these various exposures, measured in terms of deaths, years of life lost (premature deaths) or overall disease burden.

The sample table below shows the mortality tobacco-attributable fractions (expressed as percentages) for chronic obstructive pulmonary disease (COPD) classified by age, sex and sub-region. As an example of the interpretation of the information presented in Table 11, in the sub-region AFR-D (which includes such countries as Algeria and Angola) the proportion of male COPD (Global Burden of Disease code 112) mortality in the age group 30-44 which is attributable to the consumption of tobacco is 31 per cent, while for women in that age group the proportion is ten per cent.

Table 11 – Sample of the epidemiological information presented in Ezzati *et al* (2004)

Risk factor: **Tobacco**
Chronic obstructive pulmonary disease
GBD disease code: **112**

a	*0–4*		*5–14*		*15–29*		*30–44*		*45–59*		*60–69*		*70–79*		*≥80*		*Total*		
Sub-region	*M*	*F*	*M*	*F*	*M*	*F*	*M*	*F*	*M*	*F*	*M*	*F*	*M*	*F*	*M*	*F*	*M*	*F*	*All*
AFR-D	NA*	NA	NA	NA	NA	NA	31	10	40	13	39	11	35	8	36	6	36	9	23
AFR-E	NA	NA	NA	NA	NA	NA	55	34	52	31	48	27	43	21	42	23	46	24	36
AMR-A	NA	NA	NA	NA	NA	NA	80	76	80	79	82	84	80	85	80	83	80	83	82
AMR-B	NA	NA	NA	NA	NA	NA	71	46	67	44	65	45	61	42	61	49	62	45	55
AMR-D	NA	NA	NA	NA	NA	NA	9	0	28	14	29	15	26	6	25	4	22	5	14
EMR-B	NA	NA	NA	NA	NA	NA	66	68	63	43	66	30	64	29	65	20	64	29	50
EMR-D	NA	NA	NA	NA	NA	NA	75	55	59	20	55	18	51	13	51	0	54	15	36
EUR-A	NA	NA	NA	NA	NA	NA	85	65	83	60	81	54	79	55	77	53	79	54	69
EUR-B	NA	NA	NA	NA	NA	NA	89	63	87	55	83	48	74	44	63	36	76	43	63
EUR-C	NA	NA	NA	NA	NA	NA	90	43	89	42	86	37	80	43	64	25	81	35	65
SEAR-B	NA	NA	NA	NA	NA	NA	82	1	75	24	71	21	68	7	68	0	69	10	47
SEAR-D	NA	NA	NA	NA	NA	NA	74	6	68	26	66	28	61	14	59	4	65	19	44
WPR-A	NA	NA	NA	NA	NA	NA	67	39	62	43	68	45	76	59	81	75	77	68	74
WPR-B	NA	NA	NA	NA	NA	NA	48	13	19	8	20	11	37	11	39	9	33	10	20
World	NA	NA	NA	NA	NA	NA	69	19	59	25	50	23	52	19	52	18	52	20	36

* Note: NA indicates “not available”.

Appendix B – Conditions attributable to substance abuse, classified by substance (source: Ridolfo and Stevenson, 2001)

Alcohol	Tobacco	Illicit drugs
Oropharyngeal cancer	Oropharyngeal cancer	Opiate dependence
Oesophageal cancer	Oesophageal cancer	Opiate abuse
Liver cancer	Stomach cancer	Accidental opiate poisoning
Laryngeal cancer	Anal cancer	Antepartum haem. due to opiates
Female breast cancer	Pancreatic cancer	Low birthweight due to opiates
Alcoholic psychosis	Laryngeal cancer	Cannabis dependence
Alcohol dependence/abuse	Lung cancer	Cannabis abuse
Alcoholic liver cirrhosis	Endometrial cancer	Amphetamine dependence
Road injuries	Cervical cancer	Amphetamine abuse
Epilepsy	Vulvar cancer	Cocaine dependence
Alcoholic poly-neuropathy	Penile cancer	Cocaine abuse
Hypertension	Bladder cancer	Accidental poison by psychostimulants
Ischaemic heart disease	Renal parenchymal cancer	Hallucinogen dependence
Alcoholic cardiomyopathy	Renal pelvic cancer	Hallucinogen abuse
Supraventricular cardiac dysrhythmias	Respiratory carcinoma in situ	Accidental) poisoning by hallucinogens
Heart failure	Ischaemic heart disease	Antepartum haemorrhage due to cocaine
Stroke - haemorrhagic	Chronic obstructive pulmonary disease	Low birthweight due to cocaine
Stroke - ischaemic	Tobacco abuse	Hepatitis B
Oesophageal varices	Parkinson's disease	Hepatitis non A, and B
Gastro-oesophageal haemorrhage	Pulmonary circulation disease	AIDS
Alcoholic gastritis	Cardiac dysrhythmias	Infective endocarditis
Unspecified liver cirrhosis	Heart failure	Drug psychoses
Cholelithiasis	Stroke	Maternal drug dependence
Pancreatitis - acute	Atherosclerosis	Newborn drug toxicity
Pancreatitis - chronic	Pneumonia	Road injuries
Low birthweight	Peptic ulcer	Suicide
Psoriasis	Crohn's disease	
Alcoholic beverage & other EtOH poisoning	Ulcerative colitis	
Other ethanol and methanol poisoning	Ectopic pregnancy	
Fall injuries	Spontaneous abortion	
Fire injuries	Antepartum haemorrhage	
Drowning	Hypertension in pregnancy	
Aspiration	Low birthweight	
Occupational and machine injuries	Premature membrane rupture	
Suicide and self-inflicted injury	SIDS (and smoking during pregnancy)	
Assault	Fire injuries	
Child abuse	Asthma (under 15 years)	
	Lower respiratory illness (under 18 months)	
	SIDS (and post natal smoking)	
	Lung cancer (passive)	
	Ischaemic heart disease (passive)	

Note: See Codde (2002) for the ICD-9 and ICD-10 codes for the above conditions

Appendix C – Estimating the social costs of drug-attributable crime

As indicated in the main body of this report, there have been recent developments in the estimation of the social costs of drug-attributable crime which were not available when the aggregate cost *Guidelines* (Single *et al*, 2003) were produced. This Appendix discusses, in much greater detail than presented in the original *Guidelines*, theoretical and methodological issues in the estimation of these costs, as the basis for the discussion in this report of the estimation of *avoidable* drug-attributable crime costs.

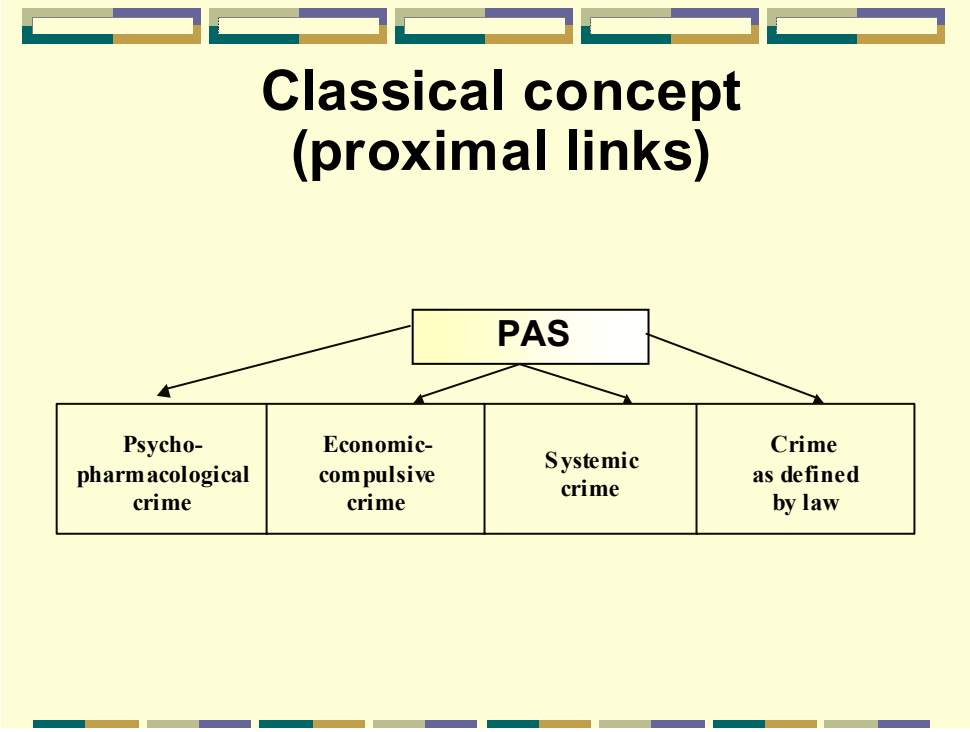
Alcohol and illegal drugs are the psychoactive substances (PAS) most commonly linked to criminal activity (although, in some jurisdictions, tobacco-related smuggling and tax evasion are important issues). However, the drug-crime link is complex and varies in terms of the persons involved, the substances consumed, and the offences committed. In order to identify and understand the avoidable costs of alcohol and drug-related crime, it is important to fully grasp the motives that ultimately drive PAS users to commit criminal acts.

The phrase "drug-related crime" is a somewhat vague and confusing expression. One way to bypass such fuzzy concepts is to look at key conceptual models. The models reviewed here are based on the results of empirical research, which gives them a measure of explanatory validity.

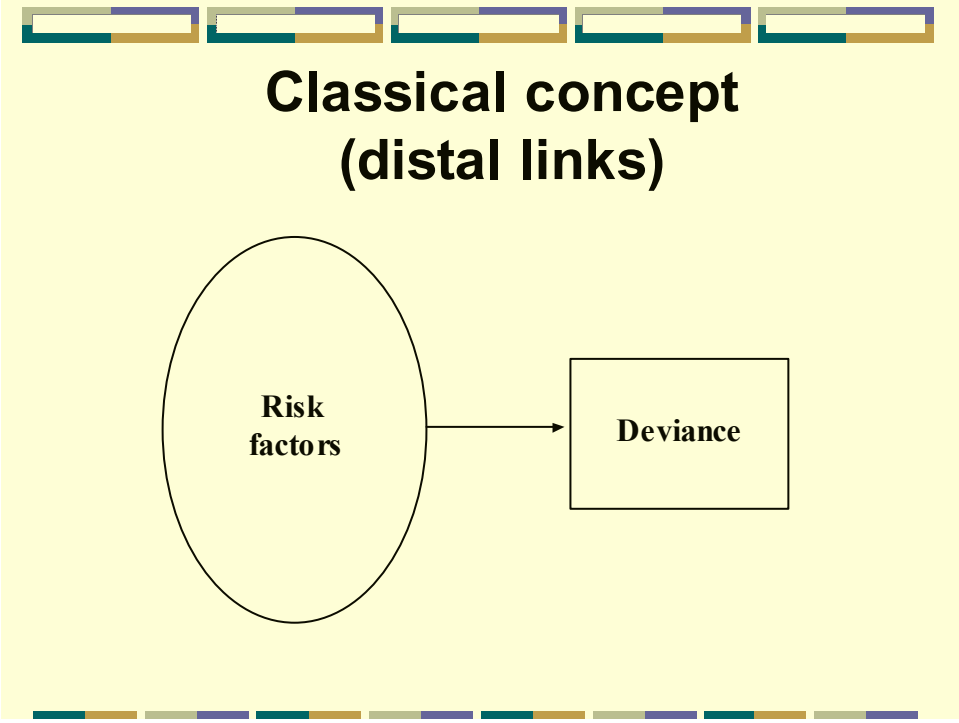
An analysis of the scientific literature reveals two major concepts or explanations for the links between substance abuse and crime. The first type of explanation—by far the most "classical" and fully developed—describes the PAS-crime link in static terms. One model that falls in this category looks at *proximal* links. This is the tripartite model developed by Goldstein (1985). It is based on different facets of the PAS-crime link, namely:

1. the psychopharmacological aspect or intoxication;
2. the economic-compulsive aspect; and
3. the systemic aspect, which refers to the distribution of illegal drugs.

The integration of these three facets of the proximal drugs-crime relationship into a tripartite explanation is based on numerous empirical studies conducted in North America and Europe. Another element should be added to this tripartite model, that of "drug crimes" as defined by law. This category of offences includes possession, trafficking, importation, as well as driving under the influence of otherwise legal substances.



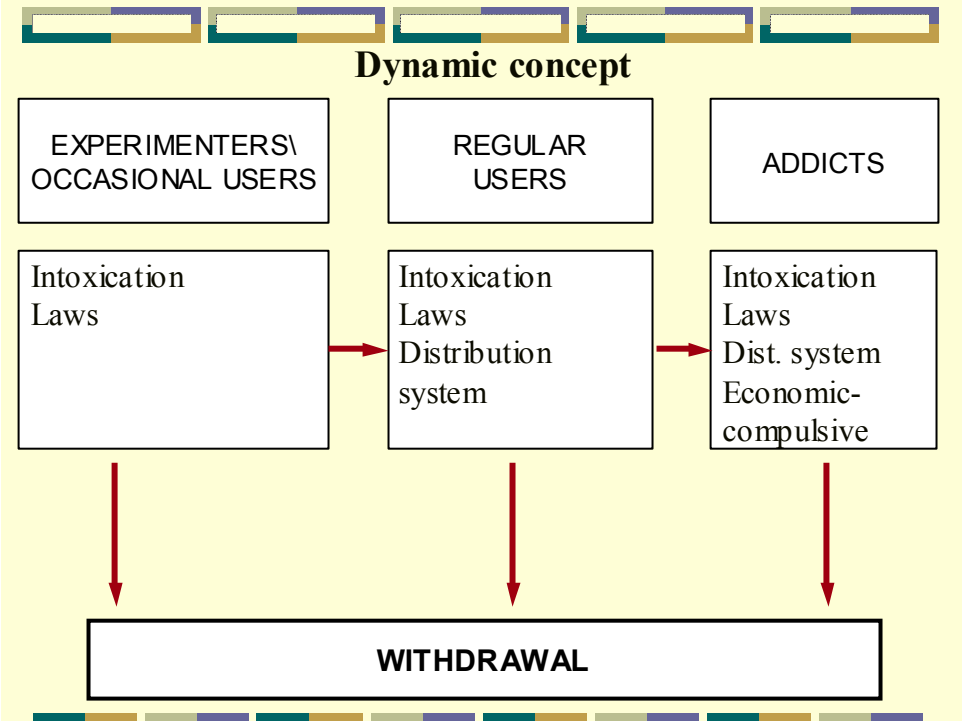
A second model looks at the *distal* links that connect both drug abuse and criminality to a range of biopsychosocial factors, commonly referred to as risk factors. We shall refer to this model as the biopsychosocial model.



The second major concept views the drug-crime link in terms of a dynamic association within a deviant trajectory or career. Generally speaking, users move through various phases, the most common being:

- initiation or experimental use,
- regular consumption without dependence,
- abuse or dependence, and
- cessation.

The PAS-crime link changes as the person moves through these successive phases. However, in seeking to explain the links between PAS and crime within each of these phases, the proponents of this concept revert to the more classical static models mentioned earlier.

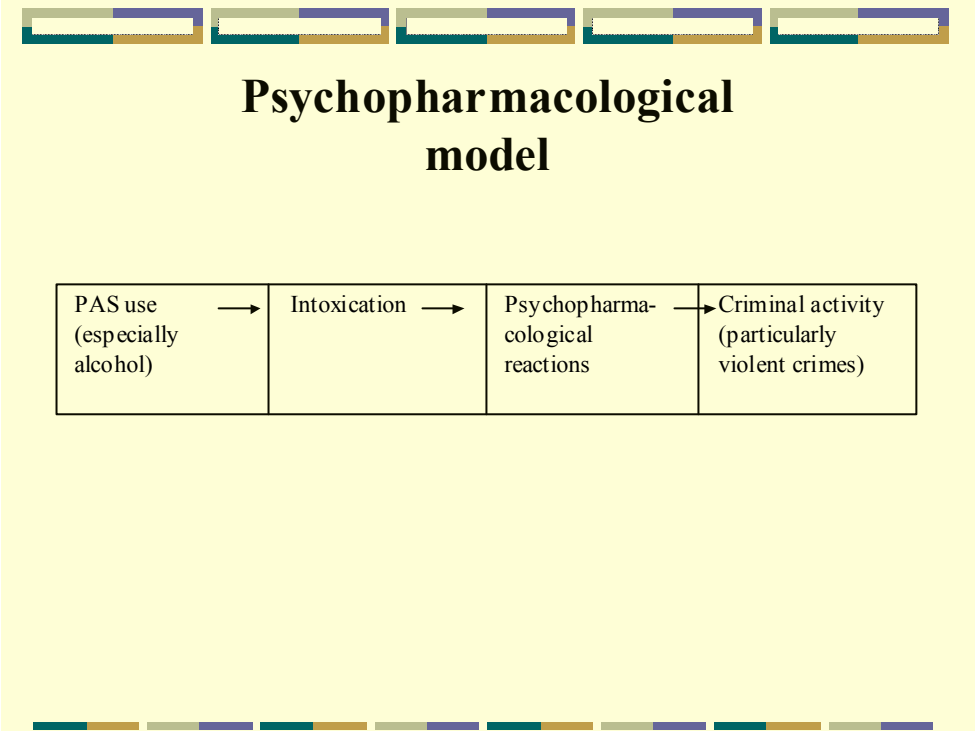


Although this concept more accurately reflects the reality of PAS users, it contributes little to the present exercise, except to remind that use—even the use of illegal drugs—is not always synonymous with dependence. This document, therefore, examines the two major classical models (both linked to the first concept), in order to gain a better understanding of the various links between alcohol/illegal drugs and criminality. It describes and analyses each facet of the drug-crime relationship, based on these two models, as well as define actions that may help to reduce drug-related crime, as well as certain costs associated with such crime.

Proximal links

a) Related to intoxication

A substance that acts on the central nervous system can no doubt play a determining role in the commission of a crime. In fact, certain PAS (usually stimulants, hallucinogens and, especially, alcohol) are commonly thought to induce various forms of aggression. The intoxicating, criminogenic action of alcohol is usually linked to disinhibition. According to this model, a combination of psychological and pharmacological factors may cause a person to behave "abnormally" and to give free rein to impulses that would otherwise be reasonably well controlled. In this psychopharmacological hypothesis, intoxication is a determining contributory factor in the commission of offences that the perpetrators would not have committed had they been sober. A variant of this model is that intoxication can serve an underlying instrumental purpose: namely to give an individual the courage to commit a crime that has already been planned, or to calm jittery nerves. In the first version of this model, intoxication leads to crimes that would not have taken place without the influence of PAS. In the second version, PAS are a tool (in the same way as a weapon or a disguise) to achieve a premeditated goal. In reality, it is difficult to establish an exact causal relationship without relying on the word of the social actor; nonetheless, it is clear in both cases that psychoactive substances play a significant contributory role.

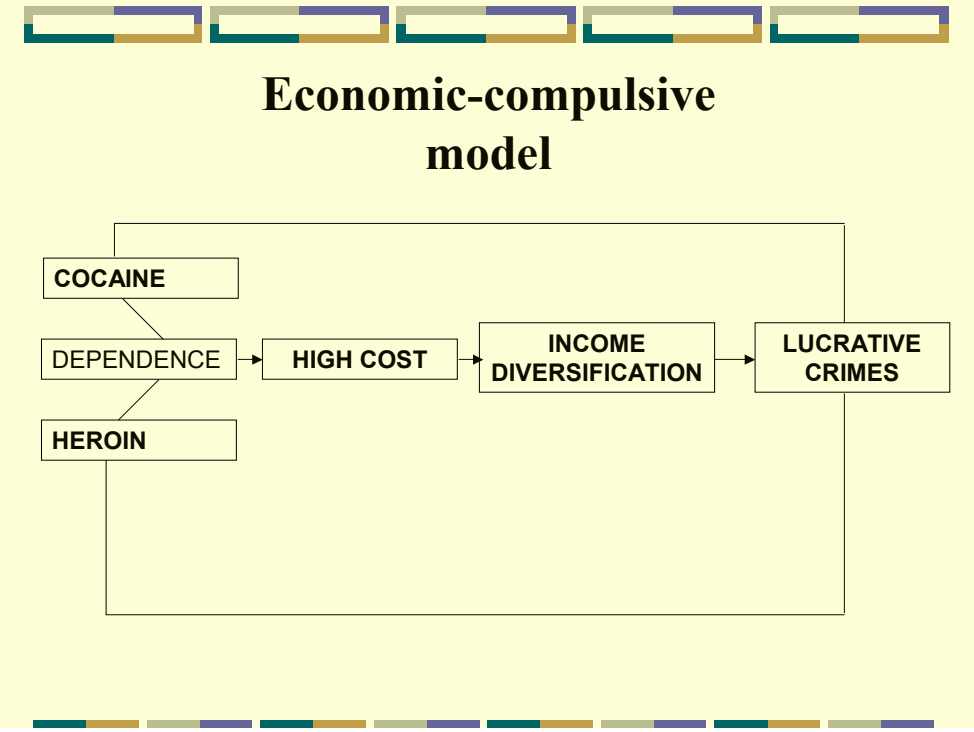


In terms of illegal drugs, there is little information on the level of consumption that is likely to lead to problems. According to the prohibitionist philosophy, all consumption entails a risk. For alcohol, which is no longer subject to the prohibitionist philosophy that prevailed in North America in the early part of the 20th century, it is possible to establish

levels of intoxication that are best not exceeded. In Canada, for example, it is illegal to drive a car when one's blood alcohol level exceeds 80 mg/100 mL. While such measurement is somewhat idiosyncratic, it is generally accepted that having more than five drinks on any given occasion is problematic and carries a risk.

b) Related to dependency

The most significant link between drugs and crime is the economic aspect associated with the purchase of illegal substances. Certain drugs, particularly heroin and cocaine are habit-forming for some users. A user who becomes dependent on one of these drugs must have it several times a day in order to avoid physiological or psychological withdrawal symptoms. Over time, use of these substances becomes extremely onerous. The crimes committed by some users who are no longer able to control their consumption can be explained, at least in part, by their need to obtain money to buy the drugs to which they have become addicted.



c) Related to the distribution system for illegal PAS

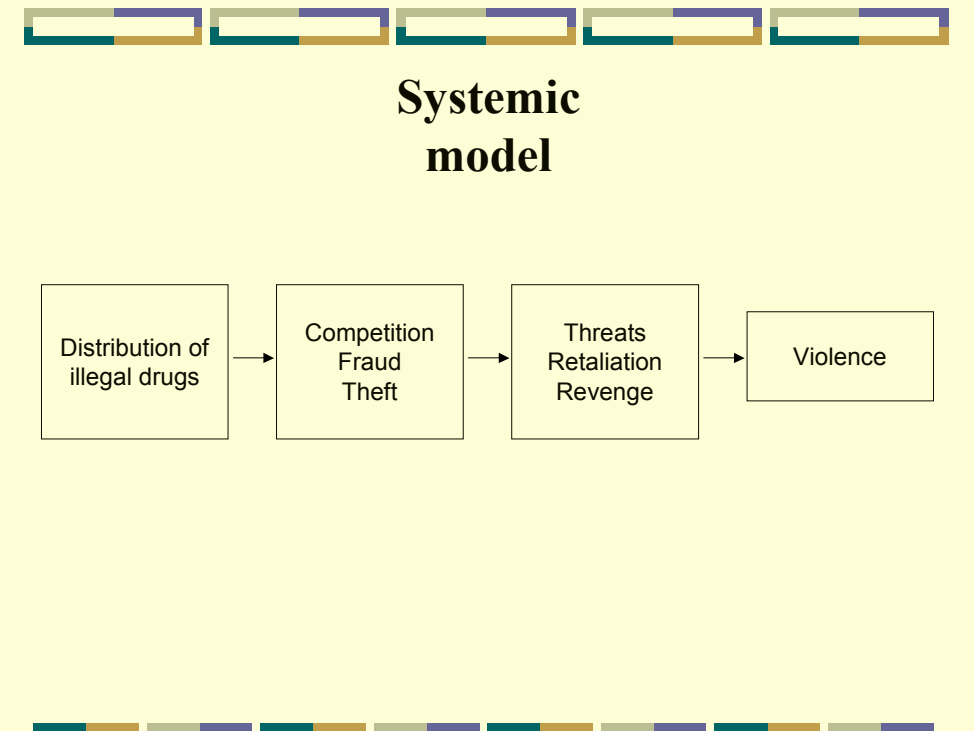
The Single Convention on Narcotic Drugs (1961, modified in 1972) and the Convention Against Illicit Traffic in Narcotic Drugs and Psychotropic Substances (1988) are designed to limit, through criminalization, the growth, production, trafficking, distribution, possession and consumption of certain drugs. Taking their lead from the United States, many countries have used these conventions as the basis for a "war on drugs." Alongside this repressive approach, a distribution system for illegal drugs has taken shape. Crimes related to the distribution of illegal drugs are generally committed in the course of selling drugs or collecting drug debts, or as a result of "turf wars" linked to

traffic in illegal drugs. These are not offences related to the intrinsic properties of drugs, but rather crimes associated with the illegal context of distribution of a product. These crimes are termed "systemic," in that they occur within an illegal commercial system.

This system encourages criminal activity in two key ways. First, the illicit nature of the drug culture promotes:

- experimentation and consumption of new products by users, sometimes leading to dependence (this is sometimes referred to as the gateway effect);
- reimbursement of accumulated debt through illegal means (resale of drugs, prostitution, and other lucrative criminal activities).

As well, this environment tends to be rife with territorial disputes between rival dealers or between dealers and dissatisfied clients. Violence is frequently used as a form of "personnel management" and retaliation takes the place of the penal justice system. It is safe to assume that individuals already drawn to violence may be strongly attracted to the methods employed in this milieu, where their skills and physical strength can be exchanged for significant monetary rewards. However, it is important to note that the violence associated with the drug distribution system is far more prevalent in large North American cities than in major European centres, which suggests that the socio-political environment plays a role in such violence.

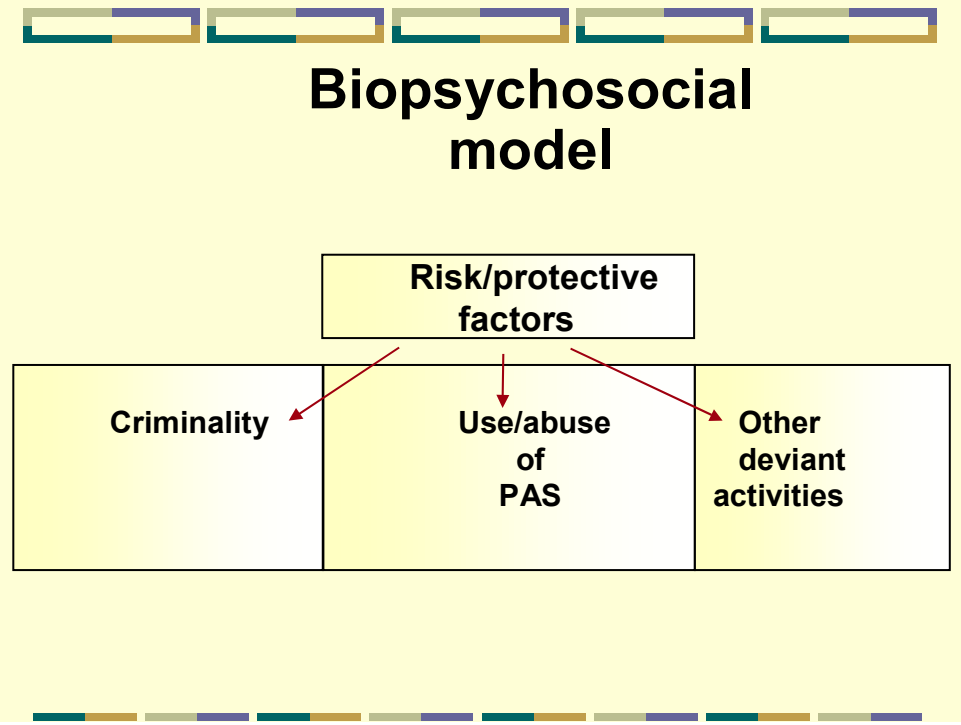


d) Defined by law

A fourth element should be added to the three elements defined in Goldstein's tripartite model, in order to adopt a broader perspective when defining drug-related crimes. Like "systemic" criminality, these types of crimes are not directly linked to the properties of illegal substances, but rather are defined under the major international conventions mentioned earlier, as well as by the laws that govern alcohol consumption and the behaviours associated with such consumption on a social level. Many PAS-related infractions, such as possession or consumption, growing or manufacturing, as well as smuggling and trafficking, fall into the category of crimes defined under law. The prevalence of such crimes, as reported in official statistics, depends far more on the means of repression in place than on actual illegal activities. Generally speaking, these statistics are greatly influenced by one type of crime (possession) and one product (cannabis).

Distal links: the biopsychosocial model

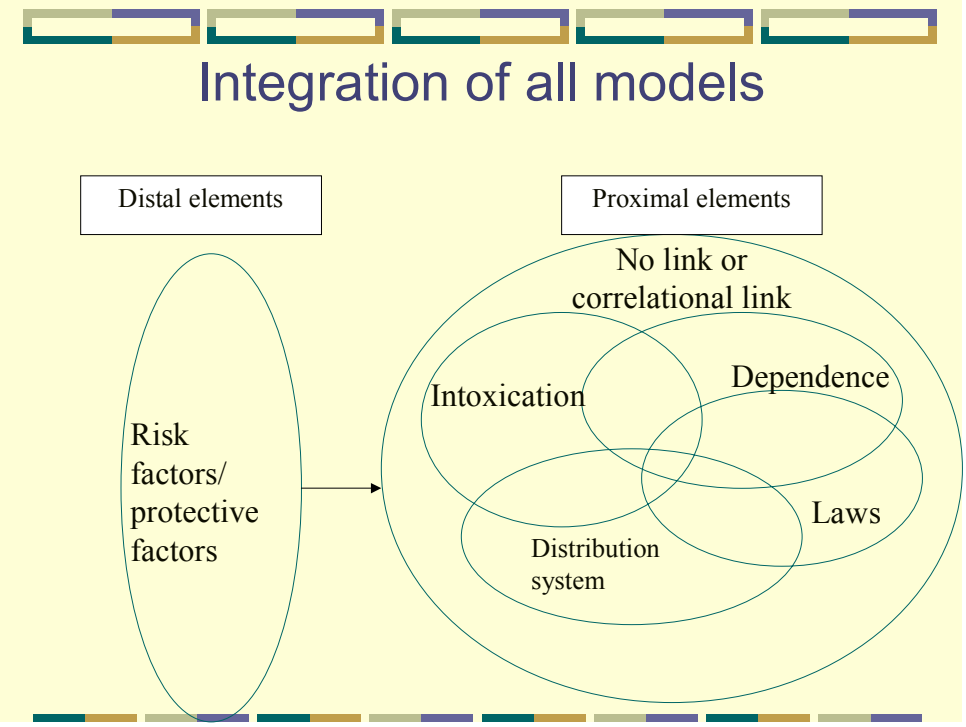
Numerous studies suggest that criminality, like drug abuse, is unequally distributed across the population. In fact, a small number of adolescents account for a large proportion of all deviant behaviour. This structural marginality is thought by some to be associated with a general syndrome of deviance. According to this concept, delinquency, drug use, as well as certain other deviant or marginal behaviours, such as early (and frequently unprotected) sexual experimentation, dangerous driving, risk-taking behaviours, etc. are linked to the presence of risk factors in the social actor's past. These risk factors (socio-demographic background, current environment, family, peers, estrangement from social institutions and socially accepted norms) "predispose" individuals to adopt a lifestyle in which intoxication, drunk driving, addiction, and crime are part of everyday life. However, a number of protective factors are thought to play a significant role in what researchers and workers in the field refer to as "resilience." In short, the studies informed by the biopsychosocial model clearly demonstrate how difficult it can be to establish exclusive causal links between psychoactive substances and crime, since the relationship is also influenced by distal links.



A calculation error to avoid

It is to be expected that the forms of crime defined earlier will overlap. Such overlapping can be significant and may result in the same crime being counted several times. For example, some individuals who commit a crime under the influence of drugs may do so to obtain drugs for their own consumption. If, in addition, the crime constitutes an offence under the law (e.g., trafficking), there is a chance that the same criminal act will be counted three times. Such double, triple or quadruple counting may give the impression that a larger proportion of crimes are drug-related than is actually the case; it may also result in overestimation of the costs that could be avoided if new measures were implemented.

The following figure illustrates the integration of the various models discussed above



Methodology of cost estimation

The basic methodology for the estimation of drug-attributable crime costs is to determine the total costs of a particular activity (for example policing or incarceration) and then to estimate the proportion of these costs causally attributable (as opposed to related) to drug use. Thus the fundamental data needs for drug-attributable crime cost studies are

- Aggregate cost data, and
- Attributable fractions.

An Australian example of how such attributable fractions may be produced is presented in Collins and Lapsley (2002) and is used here for illustrative purposes.

Attributable fractions were developed for the Australian study by the Australian Institute of Criminology (AIC) and their derivation is fully explained in Appendices C and D of Collins and Lapsley (2002). The AIC study drew on the work of Pernanen *et* al (2002) studying Canadian offenders.

Attributable fractions for prisoners were derived from the AIC DUCO (Drug Use Careers of Offenders) survey data. The analogous fractions for police detainees were derived from the AIC DUMA (Drug Use Monitoring in Australia) survey data. DUCO examines the lifetime offending and drug use careers of adult sentenced male inmates. The DUMA

collection provides illicit drug use information on people who are detained and brought to a police station.

The Australian cost study, heavily dependent on the analysis presented above, illustrates how drug-attributable fractions for crime may be derived. The type of information which can be derived from surveys such as DUCO and DUMA is illustrated in the following two tables. Table 12 presents estimates of the Australian prisoner attributable fractions (derived from DUCO data) and Table 13 presents estimates of the attributable fractions for police detainees (from DUMA).

Table 12 – Crime-attributable fractions (prisoners), by category of crime, Australia, 2001

	Violent %	Property %	Drug Offences %	Traffic Offences %	Breaches %	Disorder %	Drink Driving %	Other %
Illicit drugs only	10.8	23.4	100.0	8.4	15.2	6.3	0.0	15.9
Alcohol only	11.0	4.1	0.0	12.8	12.7	12.6	100.0	11.4
Alcohol and illicit drugs	12.6	9.4	0.0	6.8	10.8	6.3	0.0	17.4
Neither	65.5	63.1	0.0	72.0	61.4	74.8	0.0	55.3
Total drugs	34.5	36.9	100.0	28.0	38.6	25.2	100.0	44.7

Source: Australian Institute of Criminology (see Appendix C of Collins and Lapsley, 2002).

Table 13 – Crime-attributable fractions (police detainees) by category of crime, Australia, 2001

	Violent %	Property %	Drug Offences %	Traffic Offences %	Breaches %	Disorder %	Drink Driving %	Other %
Illicit drugs only	27	43	100	17	16	9	0	8
Alcohol only	7	2	0	2	5	15	100	4
Alcohol and illicit drugs	3	1	0	0	0	0	0	2
Neither	63	54	0	81	79	76	0	86
Total drugs	37	46	100	19	21	24	100	14

Source: Australian Institute of Criminology (see Appendix D of Collins and Lapsley, 2002).

To illustrate the meaning of these fractions, consider Table 12. Of all violent offences for which prisoners are incarcerated, 10.8 per cent are estimated to be causally attributable to the consumption of illicit drugs and 11.0 per cent attributable to alcohol, with drugs in total explaining 34.5 per cent of violent crime. An analytical complication thrown up by this table is that, in the Australian study, some component of crime is causally attributable jointly to alcohol and illicit drugs (in the case of violent crime 12.6 per cent). It was not possible meaningfully to disaggregate these joint fractions back to the individual drugs.

In this table, all drug offences are assumed to be attributable to drugs (the fraction is 100 per cent) and all drink-driving is assumed attributable to alcohol. Collins and Lapsley (2002) discuss in considerable detail some of the analytical and interpretational problems which can arise in producing attributable fractions from surveys like DUCO and DUMA.

Types of drug-attributable crime costs

Drug-attributable crime can lead to the imposition of a range of social costs. The major cost categories are now considered, using as an illustration the above-mentioned Australian study.

Law enforcement

Most countries will have estimates of total public expenditures on law enforcement (policing) for all types of crime, whether drug-attributable or not. These expenditures can then be allocated to the individual types of crime according to the proportions of detainee hours in police custody classified by most serious offence of detainee. In the Australian study, these data were derived from the Australian Institute of Criminology National Police Custody Survey August 1995 (see Carcach and McDonald, 1997). Appropriate proportions of these expenditures, classified according to type of crime, were then assigned to types of drug-attributable crime according to the DUMA (detainee) attributable fractions.

Criminal courts

Public expenditures on criminal courts, again, will be available in most countries. In the Australian study, they were allocated to the individual types of crime according to the proportions of police detainees classified by their most serious offence, data derived from the National Police Custody Survey (Carcach and McDonald, 1997). They were then allocated to drug-attributable crime according to the DUMA (detainee) attributable fractions.

Prisons

In the Australian study, aggregate expenditures on incarceration were allocated to the individual types of crime on the basis of data from the National Prisoner Census presented in the Australian Bureau of Statistics publication *Prisoners in Australia* and to drug-attributable crime according to the DUCO (prisoner) attributable fractions.

Customs

Customs services usually have a variety of simultaneous functions – border protection, immigration controls, prevention of smuggling, quarantine requirements and prevention of import or export of illicit drugs. In practice there would appear to be great problems in allocating joint costs between these various functions.

Organised crime

Many countries have organisations specifically dedicated to dealing with "organised crime." As was the case in the Australian study, disaggregation of the costs of running these organisations into individual components of drug-attributable crime cost can be an extremely challenging exercise.

Forgone productivity of criminals

If prisoners had not been incarcerated their labour would have been released for productive use. However, there is reason to suspect that such labour would not in all cases have been put to productive use. In the Australian study, using data from the National Prisoner Census it was possible to estimate the value in a free market of the potential output of prisoners if they were not currently incarcerated. Since there were no data available on the number of people engaged in drug-attributable crime but not detained or imprisoned, it was not possible to estimate the potential value of their labour in productive employment.

Property theft

Clearly a considerable amount of property theft is attributable to the consumption of alcohol or illicit drugs. However, conventional economic literature asserts that this type of theft does not represent a real loss to the community as a whole. Rather, as long as the property is not subsequently damaged or destroyed, it represents a redistribution of assets from the victims (or perhaps insurance company customers and shareholders) to the thieves and their customers. A problem with this analysis is that, in the process of theft and resale, a significant proportion of the property value is lost. The value of the stolen property to the thief (in terms of its resale value) is, in almost all cases, less than its value had been to the victim of the crime. The difference between the two values represents a cost to the community as a whole. While the stolen property may remain undamaged, and so physically unchanged, its value to the community as a whole has still declined.

Thus, estimates need to be made of:

- The value of drug-attributable property theft; and
- The proportion by which this value declines as a result of the theft. In the Australian case, the decline in value was estimated to be, on average, about forty per cent of the value when new of the stolen property.

Violence

Information here is needed on the health care and welfare costs of violence and the proportion of violence which is estimated to be drug-attributable.

Money laundering

Money laundering has complex economic effects which are extremely difficult to analyse and cost. Such analysis would normally be beyond the scope of studies of the social costs of substance abuse.

Legal expenses

Costs are incurred in the employment of the legal profession in crime-related cases, for example in providing defence services to accused. Such data are likely to be most difficult to obtain.

Under-reporting of crime

It can be asserted with a high degree of confidence that most estimates of the social costs of drug-attributable crime are underestimates of the "true" costs of such crime. The major reason for this confident assertion is evidence that much crime is not reported to the police.

For example, in Australia Bryant and Williams (2000) have concluded that only about 30 per cent of alcohol- or other drug-related violence was reported to the police. Carcach and Grant (2000) reported data from the 1998 National Crime and Safety Survey (Australian Bureau of Statistics) which showed that, respectively 74 per cent and 30 per cent of (most recent) incidents of household and personal offences were reported to police.

Appendix D – Evidence on the effectiveness of interventions to reduce drug-attributable crime

This Appendix surveys the evidence on the effectiveness of programs designed to reduce the extent of drug-attributable crime. It uses, as a framework, the analysis presented in Appendix E on the estimation of drug-attributable crime costs.

Proximal links

a) Related to intoxication

Appropriate interventions could have a positive impact on reducing intoxication and related crimes. Two major types of programs are available:

1. drug prevention programs of the type provided in school settings, and
2. treatment programs for persons charged with driving under the influence.

1. Drug prevention programs

Studies dealing with drug prevention programs generally conclude that such measures are effective in reducing drug use among young people (Botvin, 1990; Botvin, Schinke and Orlandi, 1995; Ennett, Tobler, Ringwalt and Flewelling, 1994; Hansen, 1992; Tobler, 1992; Tobbler and Straton, 1997). However, few studies—and still fewer meta-analyses—have looked at the impact of evaluated programs in reducing drug-related crime.

A study by Werch, Pappas and Castellon-Vogel (1996) targeted 408 prevention programs in colleges and universities in 49 American States in order to develop a profile of U.S. efforts in this area. A questionnaire was sent to program coordinators in order to gather information on the nature of the programs, as well as their impact, integration and environmental supports. Most questionnaires were completed, such that a response rate of 82% was achieved. However, the results indicated that only 34% of the programs had carried out impact assessments. Among the impacts noted, 36.4% of the 336 responding institutions reported a reduction in alcohol-related crime, and 37.2% reported a reduction in illegal drug-related crime. However, it is impossible to determine the scale of this reduction in PAS-related criminal activity.

Wilson, Gottfredson and Najaka (2001) conducted a meta-analysis that dealt with the impact of school-based prevention programs on problem behaviours in students. Only programs that: (1) were evaluated; (2) were delivered in a school setting; (3) included a comparison group; and (4) measured a variable of interest (delinquency, PAS use, drop-out rates, or antisocial behaviour) were included in the meta-analysis. A total of 165 studies (219 documents) were surveyed following research in classical databases and contact with researchers. The vast majority of these studies (80%) had been published in peer-reviewed scientific journals. The impact of programs was calculated by comparing the average effect size with the comparison group. In general, the results indicated that the prevention programs had had a weak positive effect in terms of reducing problem

behaviours (delinquency: 0.04; PAS use: 0.05; dropping out: 0.16; other problem behaviours: 0.17).

Not all prevention programs are effective. For example, traditional prevention models that rely exclusively on the transmission of information would appear to miss their target (Botvin, Botvin and Ruchlin, 1998; Ennett, 1994). Such programs overlook the fact that adolescent drug use is a multifactorial phenomenon requiring more complex preventive interventions that effectively address the factors associated with drug initiation and abuse, as well as teach concrete social skills.

2. Programs to reduce DWI offences

A meta-analysis of measures put in place to reduce driving while intoxicated(DWI) offences was carried out by Wagenaar, Zobeck, Williams and Hingson (1995). These investigators identified 125 studies that looked at any of 12 control strategies (driver's licence suspensions, fines, community service, prison terms, use of patrols and other forms of police control, etc.). Most of these studies (70%) indicated that the evaluated measures were associated with reduced rates of drunk driving and traffic accidents. On average, a reduction of 10% was achieved, which can be termed a weak positive impact for this type of crime.

b) Related to dependence

The sharp increase in the number of studies that examine the effectiveness of treatment programs for persons in the court system is largely due to a clear conclusion that has emerged in the United States - repressive measures to curb drug use are extremely onerous (Walters, 2001). Each year in the United States, some 1.5 million persons are arrested for drug-related offences (Elsner, 2005; Gilbert, 2005). Many of these arrests result in a period of detention. As a result of America's repressive policies with respect to illegal drugs (minimum sentences, prolonged sentences, reduced opportunities for parole), the number of drug addicts in correctional institutions is steadily growing. In turn, this phenomenon is increasingly causing public administrators to question current practices and to ask themselves whether incarcerating drug addicts is the best way to prevent them from re-offending.

The delivery of rehabilitation treatment, the implementation of drug replacement programs (e.g., methadone), and the medical prescription of users' drugs of choice (such as heroin) are emerging as effective means of reducing the number of criminal acts that result from the need to find money to buy drugs. Each of these measures is now analysed in order to determine their impact in reducing the rate of crimes committed under the influence of PAS.

1. Rehabilitation treatment

The *National Treatment Outcome Research Study* (NTORS) looked at a range of drug treatment programs (N=54) provided in England. Through this vast study, a total of 1,075 problem drug users were interviewed between March and July 1995. Of this number, 753 were interviewed again a year later, and 418 were interviewed at the 4-5 year follow-up point. It would appear that rates of acquisitive crime were reduced to less than one

quarter of prior rates for those who received residential forms of treatment, both after one year and after 4-5 years (Gossop, Marsden and Stewart, 2000; Gossop, Marsden, Stewart and Kidd, 2003).

Closer to home, an American study (Schildhaus, Gerstein, Brittingham, Cerbone and Dugoni, 2000) looked at pre- and post-treatment criminal behaviour in a representative sample of 1,060,000 persons admitted into drug rehabilitation programs. The results indicate that all measures related to lucrative crime (sale of drugs, theft, prostitution, etc.) were reduced by one third.[1] Paradoxically, incarceration rates increased by 17% and parole revocations increased by 26%. In seeking to explain these paradoxical effects, the authors suggest that the individuals in question may have been subject to increased surveillance once they were known to the justice system.

Although these two studies accurately portray the situation that exists in England and the United States (given the vast scale of the samples used), neither used a valid comparison group. As a result, the observed effects cannot, beyond all reasonable doubt, be attributed to the treatment provided.

A meta-analysis of 78 studies was carried out by Prendergast, Podus, Chang and Urada (2002). Each of the studies included in this meta-analysis took effects observed in clients who had received treatment and compared them with results obtained with patients who had had only minimal exposure to a rehabilitation program. The results indicate that treatment was associated with a statistically and clinically significant effect on PAS use and on the criminal activity of clients. More specifically, the average weighted effect was 0.30 for drug use and 0.13 for criminal activity. In an earlier meta-analysis, the same team (Prendergast, Podus and Chang, 2000) observed that well-established programs showed a more significant average effect; this was also the case for programs that provided a greater number of hours of exposure to treatment.

Another form of rehabilitation treatment is through drug courts. The offender population includes a large proportion of individuals with addiction problems (Brochu, in progress). Offering treatment that is specific to the needs of these individuals would provide a means of establishing contact with many addicts who are not otherwise reached by health and social services. However, offering treatment to persons in the court system requires overcoming obstacles of various kinds (Brochu and Schneeberger, 2002). Moreover, in order to ensure optimal impact on criminal behaviour in addicts, it is important to coordinate referrals to appropriate, high-quality treatment services, which is the goal Drug Courts are seeking to achieve (Belenko, 2001).

The concept of Drug Courts is not entirely new: in fact, there were isolated experiments with the idea in Chicago and New York as far back as the early 1950s. However, the number of such courts has multiplied in the United States and in Europe since the late 1980s.

[1] The same was true of DWI and weapons possession offences.

The primary objective of Drug Courts is to enable the judicial system and treatment agencies to coordinate their efforts and actively intervene with offenders who have drug problems, so as to promote abstinence (Belenko, 2001) and treatment persistence. Each actor in the judicial system (judges, lawyers, etc.) must take a short training course in addictology. The Courts require that the treatment programs they financially support provide weekly urine samples from clients and that offenders periodically account for their progress to judicial authorities. Based on the offender's progress, the judge, who adopts a supportive role, decides on the measures to be taken (incarceration, therapy, etc.).

Belenko (2001) examined 37 evaluation reports on Drug Courts in order to assess the impact of this approach. Only reports prepared by an outside evaluator were included. According to this study, 47% of offenders referred for treatment through Drug Courts complete their rehabilitation programs. Offenders referred for treatment through a Drug Court also presented a lower risk of being arrested (5.4%) and fewer days of incarceration (6.6) than comparison group subjects (21.5% and 13.6 days respectively). While these results may seem positive (Goldkamp, 2000; Goldkamp, White and Robinson, 2001; James and Sawka, 2002), it should be noted that studies measuring the impact of Drug Courts have come under considerable criticism, including criticism of a methodological order (Fisher, 2002). Therefore, one should exercise prudence in assessing the results and simply conclude that Drug Courts constitute a positive addition where a system of referral between the courts and treatment services did not previously exist.

2. Drug replacement programs

Methadone, a synthetic opioid with analgesic effects, was first used in the1960s to treat heroin addicts (Perreault, Lauzon, Mercier, Rousseau, Gagnon (2001)). The duration of action of methadone is much longer than that of heroin (use can be limited to once a day without inducing withdrawal symptoms). When the correct dosage is given, methadone induces neither euphoria nor drowsiness.

The previously described NTORS also looked at the impact of methadone maintenance programs on the criminal behaviour of program participants. It would appear that acquisitive crime was reduced to less than a third (28%) of its initial level after one year and was further reduced to 23% of the base level at the 4-5 year follow-up point (Gossop, Marsden and Stewart, 2000; Gossop, Marsden Stewart and Kidd, 2003). Once again, this study tells us nothing about the average effect size of methadone programs.

Masch (1998) conducted a meta-analysis of 24 studies that examined the impact of methadone programs on the criminal behaviour of participants. The results indicate an average effect size of 0.25 (unweighted). When the author differentiated average effect sizes by type of crime, she obtained an average effect size of 0.70 with respect to drug-related criminal behaviour and 0.23 for property crimes. Therefore, it is safe to assume that methadone maintenance programs have a significant positive effect on levels of drug-related crime.

3. Medical prescription of heroin

The prescription of injectable heroin is a controversial topic in some countries. Still, it is viewed by many as an essential means of re-establishing contact with patients who have broken with the health and social services system and are following no other form of treatment (Touati, Sueur and Lebeau, 1999).

In 1994, a new project was introduced in Switzerland that involved prescribing heroin to persons addicted to the drug. A study (Brehmer and Iten, 2001) was carried out to observe the impact of the project on the criminal behaviour of participants (1,031 persons). At the time of their admission into the program, 70% of patients reported having been involved in some form of criminal activity (primarily selling drugs, shoplifting and possession of stolen goods); 18 months after the start of therapy, this proportion had diminished to 10%. Ribeaud (2004) examined the police data concerning the same individuals. At the time of their admission into the program, approximately half of the participants had had dealings with law enforcement agencies (for offences other than heroin use or possession) during the past year. This proportion fell to 31.5% after one year of treatment and 16% during the fourth year of treatment. Incidence fell from 1.8 offences prior to treatment to 0.73 offences during the first year of treatment, followed by 0.4 offences during the fourth year. A reduction was observed for all types of crime. However, the study did not include a control group and it is impossible to determine whether this significant reduction in criminal activity is in fact attributable to the treatment provided.

The impact of medical heroin has also been studied in the Netherlands (Van den Brink, Hendriks, Blaken, Huijsman, Van Ree, 2002), this time using a more rigorous research design that included comparison groups. According to the results of this study, the number of days of illegal activity in the previous month for persons receiving injection heroin fell from 12.9 days (at the time of admission) to 2.9 days at the one-year follow-up point. Subjects in the control group presented 11.5 days of illegal activity when the study began and 8.7 days 12 months later. It can therefore be stated that the prescription of medical heroin has a significant impact on the criminal behaviour of program participants.

c) Related to the illegal distribution system for PAS

Although the American experience with alcohol prohibition would suggest that these kinds of repressive measures generate violence and that, conversely, legalization tends to appease such systemic violence, we were not able to find studies that were sufficiently rigorous from a methodological standpoint to enable us to affirm with conviction that certain measures have a definite impact on systemic criminal activity.

However, two American economists have carried out interesting analyses that warrant further discussion. The first, Resignato (2000), used drug arrest statistics and illegal drug use statistics for 23 large American cities in order correlate them with rates of violent crime in those cities. Based on the results of this study, violent crime is more strongly linked to police enforcement activities than to drug use.

The second study (Miron, 2001) looked at the links between gun control measures, drug control measures (e.g., seizure), and violent crime (e.g., homicides). Unlike the first study, which only looked at the situation in the United States, this study compared statistics from 66 countries. Here again, the results suggest that drug enforcement measures explain the differences in homicide rates reported by different countries and that the latter are in turn linked to rates of firearm possession (which are in correlation but not in a causal relationship with violence).

These two studies suggest that relaxing enforcement measures with respect to illegal drugs could have an impact on reducing rates of systemic crime. However, it is not possible to accurately estimate the potential rate of reduction.

d) Defined by law

Although the United States and the Netherlands have signed on to the same international treaties in the area of drug control (including the Single Convention on Narcotics), the two countries have adopted very different policies with respect to hard drugs. Since the mid 1970s, the American government has pursued a hard-line approach toward illegal drugs, leading to the adoption of a "zero tolerance" policy. The Netherlands, on the other hand, drew considerable disapproval in 1976 when it moved to decriminalize cannabis (Opium Act) (Korf, 2002). Since that time, many European countries have followed suit with various measures of their own. These steps are intended to isolate the various drug trades, that is to draw a distinction between the trade in cannabis (the most widespread illegal drug) and the trade in more addictive substances (such as cocaine and heroin), by decriminalizing the former. What do we know about the effectiveness of such measures?

In Canada, 4% of all criminal offences are linked to drug possession (Robinson, 2003). In view of the fact that 60% of these drug offences are in turn linked to possession of cannabis (Brochu and Cousineau, 2003), one can assume that decriminalizing cannabis possession would reduce crime rates in Canada by approximately 2%. The question then becomes whether this enforcement-related reduction would be accompanied by an increase in crimes related to intoxication.

As stated earlier, the United States has, for many years, taken a hard-line approach to illegal drug use. However, it is also important to note that 11 states decriminalized simple possession of marijuana in the 1970s without experiencing any significant increase in consumption. Moreover, in cases where an increase in consumption was observed, it was no greater than that observed in neighbouring states where no legislative change had been introduced with respect to cannabis (Single, 1981; Single 1989; Morrison and McDonald, 1995; Single, Christie and Ali, 2000).

In Australia, numerous surveys and reports have recommended the abolition of criminal penalties for possession of cannabis. In April 1987, the government of South Australia decriminalized possession of 100 grams or less of cannabis (20 grams of hashish). Researchers (Donelly, Hall and Christie, 2000) took advantage of this unique opportunity to study the effects of this liberalization, using the data from major Australian surveys on drug use to ascertain whether the new policy was having an impact on rates of self-reported drug use. They also compared reported drug use in South Australia with rates

reported in other large Australian regions. Their results indicate that lifetime use of cannabis increased more markedly in South Australia (from 26% to 32%) than in the other regions combined. However, similar increases were observed in Victoria (26% to 32%), Tasmania (21% to 33%) and New South Wales (26% to 33%), where no legislative changes had been introduced with respect to cannabis. Moreover, no differences were observed in terms of increased weekly cannabis consumption among the various regions. As a result, the researchers concluded that the differential increase in lifetime cannabis consumption could not be attributed to the policy of decriminalization introduced in South Australia.

A similar study was undertaken in 1992 following the decriminalization of cannabis in the Australian Capital Territory. In order to evaluate the impact of decriminalization, McGeorge and Aitken (1997) compared cannabis use among students at the National University of Australia with that of students at the University of Melbourne in 1992 and 1994. The lifetime consumption patterns of the two groups proved to be similar, with no observable change in either case (53.8% and 53.3% respectively). The researchers therefore concluded that decriminalization of cannabis was not associated with increased use.

All of these studies indicate that decriminalization of cannabis results in reduced rates of criminal activity as defined by law, without any concomitant increase in the proportion of users or in the number of crimes committed under the influence. It is important to note that very few cannabis users reach a level of intoxication sufficient to induce them to commit crimes they would not otherwise commit (Brochu, in progress).

Distal links: the biopsychosocial model

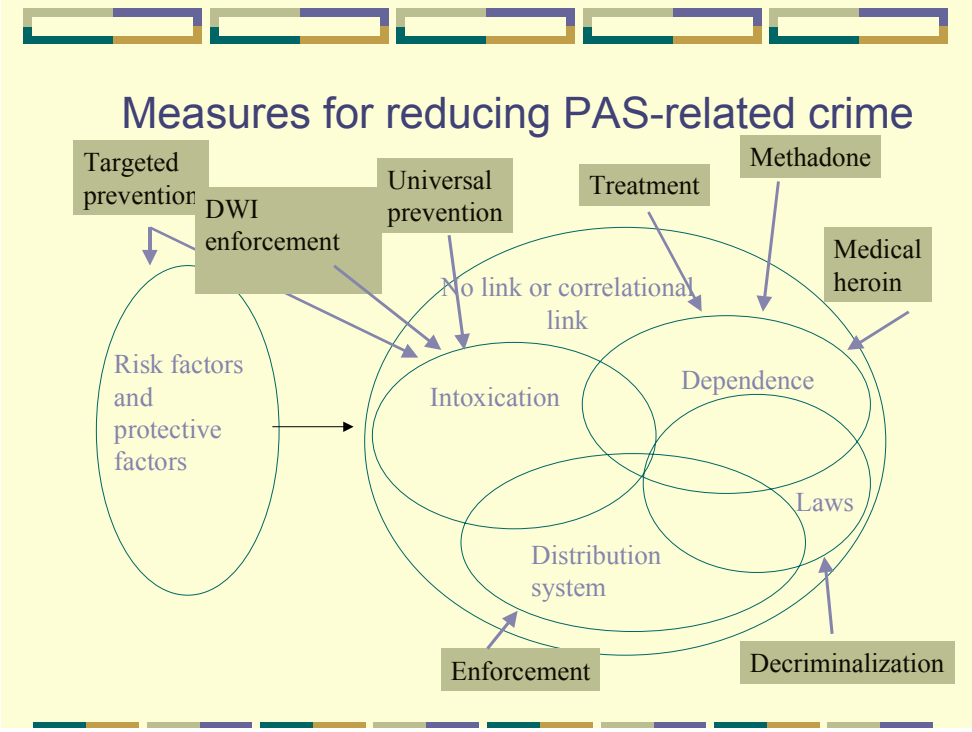
In recent years, research into risk factors has had a major impact on the development of prevention programs. Abandoning universal programs offered to all students, prevention has gone the route of targeted programs aimed at young people who have the greatest exposure to risk factors and are most in danger of falling into patterns of deviant behaviour. Such programs seek to address young people's needs by mitigating risk factors and reinforcing protective factors. Considerable hope rides on these kinds of programs as a means of substantially reducing drug abuse and criminal behaviour in the young.

In the section on the link between intoxication and criminal behaviour, the results were discussed of the meta-analysis carried out by Wilson, Gottfredson and Najaka (2001), which looked at the impact of school-based prevention programs on students' problem behaviours. While the results of this study generally showed that evaluated prevention programs had had a weak positive effect in terms of preventing problem behaviours (delinquency: 0.04; PAS use: 0.05; dropping out: 0.16; other problem behaviours: 0.17), the effect achieved was more significant when targeted prevention programs were aimed at high-risk students (0.20).

Summary

Although meta-analyses are not legion and despite the fact that further studies are needed before any definitive conclusion can be drawn, it is safe to assume that concrete measures can have a real impact on various types of drug-related crime. The measures that seem most promising encompass addiction treatment programs (including methadone programs), as well as targeted prevention programs. It would also appear that decriminalization reduces criminal activity (as defined under drug laws) without necessarily bringing about an increase in criminal activity as a result of more widespread use. Decriminalization would also reduce overall enforcement efforts with respect to drugs and might even bring about a reduction in so-called systemic crime.

The impact of such measures is not necessarily additive and it is still not possible to determine whether the application of an effective measure can indirectly influence the impact of other measures. For example, deterring individuals from criminal activity through effective, targeted prevention programs will reduce the number of individuals who commit economic-compulsive crimes. The reduction in the proportion of addicts resulting from the implementation of prevention programs could influence the effect size of rehabilitation programs by subtracting from the "treatment market" persons who would ultimately have been most receptive to therapeutic measures.



References

1. Belenko, S. (2001). Research on Drug Courts: A Critical Review 2001 Update. Columbia University, New York.

2. Botvin, J.G. (1990). Substance Abuse Prevention: Theory, Practice, and Effectiveness. *Crime and Justice, 13*, 461-519.

3. Botvin, J.G., Botvin, M.E., Ruchlin, H. (1998). School-Based Approaches to Drug Abuse Prevention: Evidence for Effectiveness and Suggestions for Determining Cost-Effectiveness. *NIDA Research Monograph*, Institute of Drug Abuse, United-States.

4. Botvin, J.G., Schinke, S., Orlandi, A.M. (1995). School-Based Health Promotion: Substance Abuse and Sexual Behavior. *Applied and Prevention Psychology, 4*, 167-184.

5. Brehmer, C., Iten, P.X. (2001). Medical Prescription of Heroin to Chronic Heroin Addicts in Switzerland - A Review. *Forensic Science International*, *121*, 23-26.

6. Brochu, S. (in preparation). *Drogue et criminalité: une relation complexe (2^{e} édition révisée).* Montréal: Presses de l'Université de Montréal.

7. Brochu, S., Cousineau, M.M. (2003). Drogues et questions criminelles : un état de la question à partir d'études québécoises. In LeBlanc, M., Ouimet, M., Szabo, D. (Eds) *Traité de criminologie empirique.* 3^{e} Édition, Presses de l'Université de Montréal, Montréal, 243-279.

8. Brochu, S., Schneeberger, P. (2002). The Effectiveness and the Difficulties of Sending the Drug Addicted Offenders to a Community Center. *Criminal Justice*, *3*, 279-287.

9. Corwyn, R.F., Benda, B.B. (2002). The Relationship Between Use of Alcohol, Other Drugs, and Crime Among Adolescents: An Argument for a Delinquency Syndrome. *Alcoholism Treatment Quarterly*, *20*, 35-49.

10. Donelly, N., Hall, W., Christie, P. (2000). The Effects of the Cannabis Expiation Notice System on the Prevalence of Cannabis Use in South Australia: Evidence from the National Drug Strategy Household Surveys 1985-1995. *Drug and Alcohol Review*, *19*, 265-269.

11. Donovan, J.E., Jessor, R., Costa, F.M. (1999). Adolescent Problem Drinking: Stability of Psychosocial and Behavioural Correlates across a Generation. *Journal of Studies on Alcohol*, *60*, 352-361.

12. Elsner, C.M. (2005). Drug Use: Extent and Correlates. In Wright & Miller. *Encyclopedia of Criminology*, 453-464.

13. Ennett, T.S., Tobler, S.M., Ringwalt, L.C., Flewelling, L.R. (1994). How Effective is Drug Abuse Resistance Education? A Meta-Analysis of Project DARE Outcome Evaluation. *American Journal of Public Health, 84*(9), 1394-1401.

14. Fisher, B. (2002). 'Doing Good with a Vengeance': Critical Reflections on 'Drug Treatment Court'. *Aquinas Symposium: "Coercion and Citizenship***,** CAMH, Toronto.

15. Gilbert, J.M.(2005). Drug Use and Criminal Behavior. In Wright & Miller. *Encyclopedia of Criminology*, 446-452.

16. Goldkamp, J.S. (2000). *What We Know About the Impact of Drug Courts: Moving Research from "Do they Work?" to "When and How They Work".* Testimony Before the Senate Judiciary Subcommittee on Youth Violence. Crime and Justice Research Institute, Philadelphia, USA.

17. Goldkamp, J.S., White, M.D., Robinson, J.B. (2001). Do Drug Courts Work? Getting Inside the Drug Court Black Box. *Journal of Drug Issues*. *31*, 27-72.

18. Gossop, M., Marsden, J., Stewart, D. (2000). Drug Selling Among Drug Misusers Before Intake to Treatment and at 1-Year Follow-Up: Results from the National Treatment Outcome Research Study. *Drug and Alcohol Review, 19*, 143-151.

19. Gossop, M., Marsden, J., Stewart, D., Kidd, T. (2003). The National Treatment Outcome Research Study (NTORS): 4-5 Year Follow-Up Results. *Addiction, 98*, 291-303.

20. Hansen, B.W. (1992). School-Based Substance Abuse Prevention: A Review of the State of the Art in Curriculum, 1980-1990. *Health Education Research, 7*(3), 403-430.

21. James, D., Sawka, E. (2002). Drug Treatment Courts: Substance Abuse Intervention Within the Justice System. *ISUMA, 3*(1), 127-133.

22. Korf, J. (2002). Dutch Coffee Shops and Trends in Cannabis Use. *Addictive Behaviors*, 27, 851-866.

23. Marsch, L.A. (1998). The Efficacy of Methadone Maintenance Interventions in Reducing Illicit Opiate Use, HIV Risk Behavior and Criminality - A Meta-Analysis. *Addiction, 93*, 515-532.

24. McGeorge, J., Aitken, C.K. (1997). Effects of Cannabis Decriminalization in the Australia Capital Territory on University Students' Patterns of Use. *Journal of Drug Issues, 27*(4), 785-793.

25. Miron, A.J. (2001). Violence, Guns, and Drugs: A Cross-Country Analysis. *Journal of Law and Economics, XLIV*, 615-633.

26. Morrison, S., McDonald, D.A. (1995). Comparison of the Social Impacts of the Legislative Options for Cannabis and Their Enforcement : An Overview of the Literature. *Australian Institute of Criminology*, Canberra, 24.

27. Newcomb, M.D. (1997). Psychosocial Predictors and Consequences of Drug Use - A Developmental Perspective within a Prospective Study. *Journal of Addictive Diseases*, *16*(1), 51-89.

28. Perreault, M., Lauzon, P., Mercier, C., Rousseau, M., Gagnon, C. (2001). Efficacité des programmes comportant l'utilisation de la méthadone. In Landry, M., Guyon, L., Brochu, S. (Eds) Impact du traitement en alcoolisme et toxicomanie. Les Presses de l'Université Laval, Laval, 95-135.

29. Prendergast, L.M., Podus, D., Chang, E. (2000). Program Factors and Treatment Outcomes in Drug Dependence Treatment: An Examination Using Meta-Analysis. *Substance Use and Misuse*, *35*(12-14), 1931-1965.

30. Prendergast, L.M., Podus, D., Chang, E., Urada, D. (2002). The Effectiveness of Drug Abuse Treatment: A Meta-Analysis of Comparison Group Studies. *Drug and Alcohol Dependence*, 67, 53-72.

31. E. Single, "The Impact of Marihuana Decriminalization", in: Y. Israel, F. B. Glazer, H. Kalant, R. E. Popham, W. Schmidt, R. G. Smart (Eds.), *Research Advances in Alcohol and Drug Problems*, Vol. VI, New York: Plenum, 198l.

32. E. Single, "The Impact of Marijuana Decriminalization: An Update", *Journal of Public Health Policy* 10:4 (1989), 456-466.

33. E. Single, P. Christie and R. Ali, "The impact of cannabis decriminalization in Australia and the United States", *Journal of Public Health Policy* 21,2 (Summer, 2000): 157-186.

34. Resignato, J.A. (2000). Violent Crime: A Function of Drug Use or Drug Enforcement? *Applied Economic, 32*, 681-688.

35. Ribeaud, D. (2004). Long-Term Impacts of the Swiss Heroin Prescription Trials on Crime of Treated Heroin Users. *Journal of Drug Issues*, 163-194.

36. Robinson, P. (2003). Adult Criminal Court Statistics, 2001/02. *Juristat, 23*(2).

37. Schildhaus, S., Gerstein, D., Brittingham, A., Cerbone, F., Dugoni, B. (2000). Services Research Outcomes Study: Overview of Drug Treatment Population and Outcomes. *Substance Use and Misuse*, *35*(12-14), 1849-1877.

38. Tobler, S.N. (1992). Drug Prevention Programs Can Work: Research Findings. *Journal of Addictive Diseases, 11*(3), 1-28.

39. Tobler, S.N., Stratton, H.H. (1997). Effectiveness of School-Based Drug Prevention Programs: A Meta-Analysis of the Research. *The Journal of Primary Prevention, 18*(1), 71-128.

40. Touati, M.A., Sueur, C., Lebeau, B. (1999). La prescription d'héroïne injectable. *Psychotropes, 5*(3), 75-89.

41. Van den Brink, W., Hendriks, M.V., Blanken, P., Huijsam, A.I., Van Ree, J. (2002). Medical Co-Description of Heroin Two Randomized Controlled Trials. Connecting Communities for Better Health Resource Center (CCBH), Utrecht.

42. Wagenaar, C.A., Zobeck, S.T., Williams, D.G., Hingson, R. (1995). Method Used in Studies of Drink-Drive Control Efforts: A Meta-Analysis of the Literature from 1960-1991. *Accident Analysis and Prevention, 27*(3), 307-316.

43. Walters, P.J., (2001). Drug Treatment in the Criminal Justice System. Taken from *www.whitehousedrugpolicy.gov*, May 8, 2005.

44. Werch, C.E., Pappas, D.M., Castellon-Vogel, E.A. (1994). Drug Use Prevention Efforts at Colleges and Universities in the United States. *Substance Use and Misuse, 31*(1), 65-80.

45. Wilson, B.D., Gottfredson, C.D., Najaka, S.S. (2001). School-Based Prevention of Problem Behaviors: A Meta-Analysis. *Journal of Quantitative Criminology,* 17(3), 247-272.

Appendix E – Evidence on the prevention of substance use risk and harm

The following tables are taken from the monograph *The Prevention of Substance Use, Risk and Harm in Australia: a review of the evidence* (Loxley *et al*, 2004a). This Monograph was commissioned by the Australian Department of Health and Ageing to provide the evidence base to underpin and inform a comprehensive prevention agenda in Australia, as part of Australia's *National Drug Strategy*. The Monograph covers evaluated Australian and relevant international approaches to the prevention of drug supply, use and harm, and approaches to prevent or delay the uptake of licit and illicit drugs by children and young people. In looking at the current application of prevention policy and strategy in Australia it identifies gaps in both prevention knowledge and effort. A summary of the Monograph appears in Loxley *et al* (2004b).

Table 14 – Summary: The effectiveness of childhood interventions

Intervention	Strength of evidence	Comment
Prior to birth		
Prevention and delaying pregnancy in young and vulnerable mothers	WFR	Little follow-up to next generation. Few studies have examined drug use
Health service reorientation (antenatal)	WFR	Universal approaches have not been studied
Family home visiting (antenatal)	**	Effects for selected population groups only
Infancy and early childhood (0-4)		
Health service reorientation	WFR	Universal approaches have not been studied
Family home visiting	**	Effects for selected population groups only
Parent education	**	Effects diminish with time
School preparation programs	**	Increasing emphasis on brain development in the first years
Primary school age (5-11)		
Family intervention	**	Some adolescent outcomes
Parent education	**	Mostly short term effects
School-based drug education	*	Need process studies. Social influences critical
School organisation and behaviour management	**	Adolescent follow-ups are being reported

Key

LI	Limited investigation		*	Evidence for implementation
ECI	Evidence is contra-indicative		**	Evidence for outcome effectiveness
WFR	Warrants further research		***	Evidence for effective dissemination
NA	Not applicable			

Table 15 – Summary: The effectiveness of interventions for young people

Intervention	Tobacco	Alcohol	Cannabis	Other Illicit	Comments
Parent education	WFR	WFR	WFR	WFR	Few programs address a single drug type
Family intervention	LI	LI	LI	*	Impacts relevant to illicit drug use can only be inferred
School-based drug education	**	*	*	*	Effects tend to be weak and short term
School organisation and behaviour management	WFR	WFR	WFR	WFR	Evidence for feasibility in high schools
Peer intervention and peer education	WFR	WFR	WFR	WFR	Little evaluation
Youth sport and recreation programs	WFR	WFR	WFR	LI	
Mentorship	LI	LI	WFR	WFR	
Community-based drug education	LI	LI	WFR	WFR	One evaluation had negative outcomes
Preventive case management	LI	LI	LI	WFR	Australian application for youth with a high number of risk factors is emerging
Community mobilisation	*	**	*	WFR	Cost-effectiveness unclear
Health service reorientation	WFR	*	WFR	LI	
Employment and training	LI	LI	LI	LI	
Social marketing	*	*	*	LI	May require delivery combined with other strategies
Law, regulation and policing	**	**	WFR	LI	Potential for wider implementation relevant to alcohol

Key

LI	Limited investigation		*	Evidence for implementation
ECI	Evidence is contra-indicative		**	Evidence for outcome effectiveness
WFR	Warrants further research		***	Evidence for effective dissemination
NA	Not applicable			

Table 16 – Summary of broad-based strategies

Broad-based interventions	Targeting:			Potential for integration with the prevention of harmful drug use
	Drug use	Risk factors	Protective factors	
Children and young people				
Early years investment		√	√	After around 15 years, could reduce some of the more severe drug use problems
School effectiveness		√	√	From 5 to 10 years, could reduce some illicit drug use. Earlier universal impacts on licit drug use
Crime prevention	√	√	√	Targets alcohol and illicit drugs. Strategies such as incarceration can increase drug problems
Homelessness strategies		√		Aggregating high-risk use could increase use
Adults				
Health promotion	√	√	√	Directly targets harmful drug use
Cardiovascular disease	√			Targets include tobacco and heavy alcohol use
Cancer	√			Targets include tobacco and heavy alcohol use
Community				
Injury	√		√	Targets include heavy alcohol use
Health education				Widely used despite having little evidence
Reducing differentials in socio-economic status		√		Evaluation of investments could advance knowledge of potential to reduce drug problems
Mental health promotion	√	√	√	Targets include harmful drug use
Community improvement			√	Evaluation could explore improvements in developmental protective factors

Key: √ indicates based on theoretical, policy or, in rare cases, an empirical basis for linking the broad policy area to reductions in harmful drug use or to modifications in developmental risk and protective factors.

Table 17 – Summary: The effectiveness of demand reduction interventions

Intervention	Tobacco	Alcohol	Illicit	Comments
Drug treatment	***	***	***	Strong evidence for nicotine replacement, alcohol treatment and methadone; less for psychostimulant or cannabis treatment
Health service reorientation				
Brief interventions	***	***	*	Highly cost effective for tobacco and alcohol
Targeted approaches for pregnant woman	**	LI	LI	Can reduce smoking, low birth weight and pre-term birth
Workplace interventions				
Drug testing in workplace	NA	WFR	WFR	No well-controlled efficacy studies
Pre-employment screening	NA	LI	LI	No well-controlled efficacy studies
Drug testing in high-risk settings	NA	*	*	Essential in very high-risk work e.g. for pilots
Brief interventions	WFR	WFR	LI	Many workplaces provide opportunities to intervene in high-risk groups e.g. young males
Community-based interventions				
Health promotion	*	*	WFR	Evidence of good acceptability within host communities
Focused on structural policy change	LI	**	LI	Target youth alcohol access, liquor and drink-driving law enforcement
Social marketing	**	*	WFR	National campaigns reduce overall smoking prevalence. Drinking behaviour change can be achieved but is difficult to sustain
National drinking guidelines and standard drink labelling	NA	*	NA	Should not be evaluated in isolation from other prevention strategies
Sub-populations: generic interventions targeting all drug types				
Treatment for co-morbid mental health and substance use issues	WFR			No direct evidence to support one form of treatment over another
Programs for indigenous people	WFR			Needs more research negotiated with indigenous community-controlled organisations
Treatment for the elderly	WFR			Not well addressed
Programs to reduce demand among the elderly	WFR			Improved screening in health care settings; preventing benzodiazepine dependence

Key

LI	Limited investigation		*	Evidence for implementation
ECI	Evidence is contra-indicative		**	Evidence for outcome effectiveness
WFR	Warrants further research		***	Evidence for effective dissemination
NA	Not applicable			

Table 18 – Summary: The effectiveness of law enforcement interventions for licit drugs

Intervention	Strength of evidence	Comment
Tobacco		
Restriction of advertising and sponsorship	**	Strong evidence that advertising controls and restrictions reduce tobacco consumption
Maintaining price disincentives	***	Strongest evidence that increases in price cause decreases in consumption
Health warning and control of pack design	LI	Limited research
Working with industry	LI	Little empirical data
Alcohol		
Restrict alcohol promotions to young people	WFR	Reasonable rationale and evidence linking exposure to ads with later drinking. Difficult area to research
Increase price through taxation to reduce consumption and harm	**	Very strong evidence-based rationale. Price increases almost invariably reduce consumption and harm
Hypothecated taxes on alcohol to fund treatment and prevention programs	***	Very strong rationale, including price increase. Controlled Australian evaluation with positive results
Outlet density	WFR	Strong rationale but no model for implementation
Outlet trading hours	**	Strong rationale; recent Australian studies have linked harms with late night trading
Responsible alcohol service and enforcement of liquor laws	** (with visible law enforcement) ECI (without enforcement)	Evidence of program effectiveness with support and appropriate law enforcement. Poor effectiveness in community-wide applications in absence of relevant law enforcement
Restrictions on price discounting	**	Very strong rationale; general relationship between price, consumption and harm. Specific evidence re Happy Hours
Licensee codes of conduct (Accords)	** (when enforced) ECI (when not enforced)	Strong rationale. Evidence for reductions in violence though results depend on external pressures for compliance e.g. enforcement of liquor laws
Dram shop laws	*	Good rationale. Evidence of deterrent effect in US and Canada
Licensing restrictions in indigenous communities	**	Most effective when part of a broad strategy and have indigenous community support
Declaration of indigenous communities as 'dry'	**	Can be effective but communities need support to enforce them. Must be under community control

Key

LI	Limited investigation		*	Evidence for implementation
ECI	Evidence is contra-indicative		**	Evidence for outcome effectiveness
WFR	Warrants further research		***	Evidence for effective dissemination
NA	Not applicable			

Table 19 – Summary: The effectiveness of law enforcement interventions for illicit drugs

Intervention	Strength of evidence	Comment
Role of law enforcement in reducing demand in the community		
Role of social norms in shaping illicit drug use (declarative)	WFR	Direct evidence is slight because of methodological difficulties but evidence from other areas of crime suggestive of effect
Role of social norms in shaping illicit drug use (general deterrence)	WFR	Sound theoretical base. One Australia study with limited support
Role of social norms in shaping illicit drug use (specific deterrence)	LI	Unsupported by the existing evidence
Role of law enforcement in reducing demand among users		
Combined targeted law enforcement and community development	*	National evaluation of US Weed and Seed program demonstrated effectiveness
Use of civil remedies to control drug and disorder problems	*	Randomised field trial of sites in California. Program effective in reducing drug offences
Police crackdowns	*	Various large and small studies in US and Australia. Evidence of effectiveness and little of displacement. Unintended negative consequences can occur
Encouraging drug users into treatment	WFR	One Sydney study found that drug users rated law enforcement as a motivator to enter treatment
Supply-side drug law enforcement		
Border protection – police	WFR	Most evaluation based on quantities seized
Border protection – customs services	WFR	Most evaluation based on quantities seized
National Heroin Signature program	WFR	Analysis continuing and results appear good but no publication
Cannabis law reform	*	Australian comprehensive review of models
Control of cultivation, manufacture and supply of illicit drugs	WFR	Evidence mainly related to quantities seized, number of charges laid and users' assessments of availability
Asset confiscation	WFR	Varies from State to State. Law Reform Commission recommends non-conviction approach

Key

LI	Limited investigation		*	Evidence for implementation
ECI	Evidence is contra-indicative		**	Evidence for outcome effectiveness
WFR	Warrants further research		***	Evidence for effective dissemination
NA	Not applicable			

Table 20 – Summary: The effectiveness of judicial procedures for the reduction of drug-attributable crime

Intervention	Strength of evidence	Comment
Diversion programs		
Diverting young offenders into early intervention services	*	Range of programs showing short term and long term gains
Diversion programs in the general community	WFR	Based on sound principles. International literature demonstrates effectiveness and improved health and well-being of participants
Courts		
Drug courts	*	Evaluations of US courts indicate effectiveness but have some methodological weaknesses. NSW evaluation demonstrates cost-effectiveness
Programs in prisons		
Drug deterrence and detection	WFR	Strongly supported by prisons. Evaluation mainly limited to counting seizures or positive tests
Differential penalties	WFR	On trial in Victoria and NSW. Good theoretical base
Provision of methadone	*	Evaluated in NSW and found to reduce injecting risk and drug use under appropriate circumstances
Dug free units	WFR	One US evaluation recommends extension to pre-release programs
Reward programs	WFR	Good theoretical base. No specific evaluations
Education	WFR	No specific evidence for effectiveness but similar programs successful in the community
Transitional support and release preparation	WFR	No direct evidence but sound theoretical base
Provision of bleach to decontaminate injecting equipment	WFR	Has the potential to kill HIV. Not known whether bleach will destroy hepatitis C (HCV)
Needle and syringe exchange	*	International experience positive

Key

LI	Limited investigation		*	Evidence for implementation
ECI	Evidence is contra-indicative		**	Evidence for outcome effectiveness
WFR	Warrants further research		***	Evidence for effective dissemination
NA	Not applicable			

Table 21 – Summary: The effectiveness of tobacco and alcohol harm reduction interventions

Intervention	Strength of evidence	Comment
Tobacco harm reduction		
Light cigarettes	LI	Likely to reduce harm; concern about misleading promotions which may cause increased consumption
Alternative nicotine delivery systems	WFR	Good theoretical base and reports of successful national implementation in some countries. No apparent evaluation data
Regulations to reduce passive smoking	***	Strongest evidence that legislation reduces environmental tobacco smoke exposure to non smokers
Alcohol harm reduction: drink-driving		
Lower BAC limits for young drivers	*	Evidence base inconsistent although half of studies show effectiveness
Random breath testing implementation	***	Strongest evidence of effectiveness Australia-wide
Ignition interlocks	**	Few but large scale studies with positive outcomes; sound rationale
Designated driver schemes	*	Sound rationale. Modest success in US studies. Successful Australian implementation although some studies have reported compliance problems
Alcohol harm reduction: other		
Thiamine supplementation to reduce Wernicke-Korsakoff's syndrome	**	Supplementation of food reduces brain damage related to heavy alcohol use. Effective Australian implementation through supplementation of flour although the most cost-effective approach has been found to be supplementing beer
Harm reduction through licensing codes of conduct	*	Evidence for short-term reductions in violence in 2/3 studies
Staggered closing times	ECI	Limited evidence. Increases harm if results in overall extension of trading hours
Plastic (or tempered) glasses	*	Sound theoretical rationale. No research evidence of effectiveness but anecdotal evidence of reductions in injuries
Food service	*	Known biological mechanism and evidence-based rationale. Not specifically evaluated
Harm reduction educational approaches	*	One well-controlled Australian study. Great potential for wide dissemination
Alcohol harm reduction: strategies in indigenous communities		
Night patrols	WFR	No outcome studies. Sound rational, wide implementation with strong community support
Sobering-up shelters	WFR	Minimal evaluations. Sound rational, wide implementation with strong community support

Key

LI	Limited investigation		*	Evidence for implementation
ECI	Evidence is contra-indicative		**	Evidence for outcome effectiveness
WFR	Warrants further research		***	Evidence for effective dissemination
NA	Not applicable			

Table 22 – Summary: The effectiveness of illicit drug harm reduction interventions

Intervention	Strength of evidence	Comment
Education to users about preventing heroin overdose	*	Sound rationale. Few evaluations
Emergency services and police protocols for overdose	*	Improvements in overdose callout rates noted
Treatment of opiate dependence to reduce risk of overdose and blood-borne viruses	***	Highest level evidence shows that engagement in treatment, especially methadone, is protective against overdose and HIV/AIDS. Evidence for protection against hepatitis C is more equivocal but rationale is sound.
Provision of naloxone for peer administration	WFR	Sound rationale. Promising international data but no comprehensive evaluations
Needle and syringe distribution	***	Highest level international and Australian evidence of efficacy including economic evaluation
Supervised injection centres	*	Sound rationale. Interim Australian evidence suggests lives saved
Hepatitis B vaccination	***	Known biological mechanism and strong rationale. Widely implemented
Retractable syringes	ECI	Weak rationale based on international reviews. Unlikely to be of benefit and may cause harm
Information campaigns for users of dance drugs	WFR	Sound rationale. No formal evaluation
Guidelines for provision of safe venues at dance parties/nightclubs	WFR	Harm is often a function of environment, such as overheating. Sound theoretical rationale for guidelines
Pill testing at venues	WFR	Sound rationale. Consumer acceptance. Widely implemented internationally. No evaluation
Pill testing at home	LI	Some concern about accuracy of tests
Dyes in benzodiazepines to reduce drink-spiking	LI	Plausible rationale but no evidence of efficacy
Harm minimisation drug education	WFR	Evidence of effectiveness for alcohol but no studies of applications to illicit drug use

Key

LI	Limited investigation		*	Evidence for implementation
ECI	Evidence is contra-indicative		**	Evidence for outcome effectiveness
WFR	Warrants further research		***	Evidence for effective dissemination
NA	Not applicable			

Appendix F – Implications of unrecorded alcohol production and consumption for the estimation of the aggregate and avoidable social costs of alcohol

Unrecorded alcohol represents a potentially significant problem for the estimation of the aggregate and therefore the avoidable costs of alcohol in many countries. There are several types of alcohol production or consumption that are often not reflected in official statistics on alcohol (Single, 2004). These include both commercially produced alcohol and non-commercial sources of production. Unrecorded commercially produced alcohol includes: (1) commercial products sold illegally; (2) legally imported commercial alcohol for personal consumption; (3) illicit importation of commercial alcohol; and (4) by-products of commercial production and commercially produced non-potable alcohol. There are also at least three sources of unrecorded non-commercial alcohol: (5) alcohol illegally produced on a large scale; (6) local small-scale production of alcohol outside of the formal economic system; and (7) home production for personal use.

The sources of unrecorded alcohol consumption vary widely between countries. For example, the legal importation of small amounts of commercial alcohol for personal use does not appear to add appreciably to alcohol consumption in many countries, but it may be significant in tourist-based economies and regions where trade barriers against alcohol imports have been eliminated (as in the EU). Illicit importation of commercial alcohol tends to be significant where there are wide discrepancies in the price of alcohol in neighboring jurisdictions. Unlicensed production contributes significantly to overall consumption in several countries such as Brazil (Vaissman, 2004), India (Gaunekar et al., 2005), Mexico (Rosovsky, 2004), Russia (Zaigraev, 2004), Ukraine (Magdenko, 2005) and Zambia (Haworth, 2004), but less so in Canada (Single and Giesbrecht, 1979).

By its very nature, unrecorded consumption and production from these various sources is hard to estimate and there are insufficient studies on unrecorded alcohol to warrant a meta-analysis of the factors underlying the amount of unrecorded consumption in a society. However, case studies of unrecorded consumption (see, e.g., Haworth and Simpson, 2004) indicate potential determinants. Clearly, high prices and limited availability of commercially produced alcohol are important influences—countries with high levels of unrecorded consumption are invariably those where commercially produced alcohol is very expensive or otherwise difficult for many consumers to obtain (Single, 2004). Other likely factors include:

- the extent to which small-scale producers can gain entry into the market; availability of natural resources and raw materials for local small-scale production;
- the availability of labour and technology for non-commercial production; public attitudes regarding non-commercial products; and
- the extent to which laws governing production are enforced (Single, 2004).

Regardless of its source, unrecorded alcohol consumption has significant economic consequences, both positive and negative (Single, 2004). On the positive side, there are

clearly much the same social benefits from unrecorded alcohol as those that derive from commercial alcohol. To the extent that unrecorded alcohol contributes to low-risk drinking, there are also the same cardio-vascular and other health benefits that result from drinking commercial alcohol in low quantities. Non-commercial production also brings economic benefits to local communities, providing employment and income to producers (often supplemental income) and lower priced alcohol to consumers.

However, the consumption of unrecorded alcohol also involves negative economic consequences, including higher risk of a variety of chronic diseases, increased risk of industrial and traffic accidents, and social problems such as unemployment, productivity losses, marital discord and alcohol-related violence. As documented in studies on the costs of alcohol use (e.g., Collins and Lapsley, 2002; Rice et al., 1990; Single et al., 1998), these adverse consequences of alcohol use have significant impacts on the economy, increasing health care and productivity costs and forcing governments to invest in prevention and research. In addition to those adverse consequences that flow from alcohol use *per se*, unrecorded consumption tends to carry additional negative consequences. Although toxicological tests indicate that noncommercial alcohol may not be as toxic as widely thought (Nuzhnyi, 2004), it is clear that unrecorded alcohol contributes to levels of alcohol poisonings in many countries. Moreover, non-commercial alcohol production generally entails a loss of government revenue and a loss of business for commercial producers, who view small-scale non-commercial production as unregulated and unfair competition.

For economists, alcohol researchers and policy makers, unrecorded alcohol represents a serious impediment to research and planning. In countries where unrecorded alcohol constitutes a significant share of consumption, it leads to underestimation of alcohol consumption and difficulties in monitoring alcohol problems. Cost estimation studies are generally restricted to the economic consequences to the legitimate market economy. It would require a complex and demanding economic framework such as general equilibrium modeling to measure the full ramifications of non-commercial alcohol in situations where it constitutes a major share of consumption. There has never been an economic analysis of non-commercial alcohol using such a framework.

The major problem arising from unrecorded consumption is the underestimation of consumption levels, which in turn results in the underestimation of alcohol-attributed mortality and morbidity. Where there is significant use of unrecorded alcohol, the prevalence of persons consuming alcohol at levels associated with a higher risk of death and disease is underestimated, thus causing a serious underestimation of the number of deaths and hospitalizations caused by alcohol use. As the health care and productivity costs largely flow from these mortality and morbidity estimates, the estimated costs can be seriously underestimated. Enforcement and prevention costs may also be disproportionately higher in societies with high rates of non-commercial alcohol use.

There are certain steps that researchers may take to address or at least ameliorate this problem. Where it is suspected that unrecorded alcohol accounts for a significant share of consumption, special studies should be undertaken on the nature and magnitude of unrecorded sources of alcohol. The resulting estimates should be incorporated in the calculations when determining levels of alcohol-attributable mortality and morbidity. It

must be recognized, however, that the problem of unrecorded alcohol consumption is largest in those countries that can least afford to conduct special studies.

References

1. Collins D., Lapsley H. Counting the cost: estimates of the social costs of drug abuse in Australia in 1998-9. Canberra: National Drug Strategy, 2002.

2. Gaunekar G., Patel V., Jacob K., Vankar G., Mohan D., Rane A., Prasad S., Johari N., Chopra A. "Drinking Patterns of Hazardous Drinkers: A Multicenter Study in India", pp. 125-144 in: Haworth, A., and Simpson, R. Moonshine Markets, New York and Hove: Brunner Routledge, 2004.

3. Haworth A. "Local Alcohol Issues in Zambia", pp. 41-66 in: Haworth, A., and Simpson, R. Moonshine Markets, New York and Hove: Brunner Routledge, 2004.

4. Haworth, A., Simpson, R. Moonshine Markets, New York and Hove: Brunner Routledge, 2004.

5. Madenko L. *Drinking Styles, Alcohol Policy and the Birth of a Nation: The first Wave of Ukrainian Independence 1991-1998*, doctoral dissertation, University of Toronto Department of Public Health Sciences, 2005.

6. Nuzhnyi V. "Chemical Composition, Toxicity and Organoleptic Properties of Noncommercial Alcohol Samples", pp. 177-200 in: Haworth, A., and Simpson, R. Moonshine Markets, New York and Hove: Brunner Routledge, 2004.

7. Rice D., Kelman S., Miller L., Dunmeyer S. The Economic Cost of Alcohol and Drug Abuse and Mental Illness, 1985 (DHHS Publication No. ADM90-1694). San Francisco: Institute for Health and Ageing, University of California, 1990.

8. Rosovsky H. "The Reporting of Alcohol Use Through Personal Diaries in Two Mexican Communities", pp. 103-124 in: Haworth, A., and Simpson, R. Moonshine Markets, New York and Hove: Brunner Routledge, 2004.

9. Single E. "Key Economic Issues Regarding Unrecorded Consumption", pp. 167-176 in: Haworth, A., and Simpson, R. Moonshine Markets, New York and Hove: Brunner Routledge, 2004.

10. Single E., Giesbrecht N. "The 16 per cent Solution and Other Mysteries Concerning the Accuracy of Alcohol Consumption Estimates Based on Sales Data." British Journal of Addiction 74: 165-178.

11. E. Single, L. Robson, X. Xie, J. Rehm, "The economic costs of alcohol, tobacco and illicit drugs in Canada, 1992", Addiction 93:7 (1998): 983-998.

12. Vaissman M. "Licit and Illicit Beverages in Brazil", pp. 87-102 in: Haworth, A., and Simpson, R. Moonshine Markets, New York and Hove: Brunner Routledge, 2004.

13. Zaigraev G. “The Russian Model of Noncommercial Alcohol Consumption”, pp. 31-40 in: Haworth, A., and Simpson, R. *Moonshine Markets*, New York and Hove: Brunner Routledge, 2004.

Appendix G – Estimating the present value of social benefits resulting from future reductions in substance abuse

Policy makers or policy advocates will often find it very useful to have an estimate of the value of the social benefits which may accrue from reductions in substance abuse, and avoidable cost estimates can form the basis for such calculations. This process is not, in itself, cost benefit analysis (CBA) since it concentrates solely on benefits but it can be readily extended into CBA.

In making estimates of the present value of the future time stream of benefits resulting from reduced substance abuse, various decisions have to be made about the impact of substance abuse and reductions to that abuse:

1. **What are the avoidable costs of abuse of the substance under study?** Estimation of this figure is discussed in the main body of this report;
2. **What reduction in substance abuse will be postulated?** For example, analysis might be undertaken of a reduction of smoking prevalence from 25 per cent to 20 per cent of the total relevant (a reduction of five percentage points), levelling out at that percentage;
3. **What proportion does this reduction represent of avoidable abuse of the substance under review, and so of its avoidable costs?** For example in the above smoking example, if the minimum achievable level of smoking prevalence were considered to be 15 per cent, the five percentage point reduction in prevalence would represent a 50 per cent reduction in avoidable smoking prevalence;
4. **What is the relationship between reductions in avoidable abuse and avoidable costs?** For example, a one-to-one relationship would mean that the above 50 per cent reduction in avoidable smoking prevalence would lead (after an appropriate time) to a 50 per cent reduction in smoking avoidable costs.
5. **Over what time period would the reduction in substance abuse occur?** Determination of this figure might be based on national or international experience in feasible rates of prevalence reduction;
6. **What path over time will the prevalence reduction take?** For example, will prevalence fall by an equal percentage point amount each year (a linear progression) or will the impact of anti-abuse policies decline over the life of these policies (a geometric progression)? The latter is often suggested by experience in public health campaigns. In anti-smoking campaigns it has been observed that, as smoking prevalence falls, it can become progressively harder to reduce the rate further as a result of delays in the reaction of more highly addicted smokers.

7. **What are the lags between falls in substance abuse and reductions in the various categories of social costs?** Reduced heroin abuse may lead to prompt reductions in the *incidence* of hepatitis but reductions in the *prevalence* of the disease will be much slower. On the other hand, a reduction in smoking prevalence will lead to a virtually simultaneous reduction in smoking-attributable fire injuries and deaths. The time paths of cost reductions are important because, given that discounting is an essential part of the process of determining the present value of reductions in substance abuse, the later cost reductions occur the lower is their present value. All other things being equal, a program which yields its benefits earlier would be the preferable program. This is because program benefits could be reinvested in other programs to yield a further rate of return and the earlier the benefits accrue the greater will be the reinvestment benefits;

8. **What is the discount rate to be chosen to convert the future time stream of cost reductions to a present value in a single year?** There are serious theoretical problems in the choice of the appropriate discount rate and, as the *Guidelines* say, there is no internationally agreed discount rate, and even in a single country economists dispute the appropriate rate. The *Guidelines* accordingly advise that studies should provide several estimates corresponding to different discount rates and that studies should include discount rates of five per cent and ten per cent among those provided, in order to facilitate comparability with studies in other countries.

9. **Over how many years are the benefits assumed to accrue?** If the reduction in substance abuse is permanent, the reduction in abuse costs will also be permanent. However, the discounting process means that, beyond a certain period, and at any discount rate significantly above zero, the present value of future benefits accruing many years ahead will be small. It is, therefore, recommended that time periods in the order of 20 years should be utilised in these calculations.

Appendix H – Issues relating to the estimation of aggregate and avoidable costs of substance abuse in Central and Southern America

The following observations are drawn from the reported experience of, and the paper presented by, Dr Augusto Pérez-Gómez and his CICAD colleagues, together with discussions at the 2005 Heath Canada Ottawa workshop. These observations reflect some of the issues relevant to the undertaking by developing countries of studies of the aggregate avoidable costs resulting from substance abuse.

There are a number of constraints in undertaking such studies which need to inform the study design, analysis and interpretation of results.

Data

Where there are no official statistics and no reliable data, it may be necessary to use proxy data to model from a comparator country. Ideally, this country should be as similar as possible to the study country. When the proxy data are epidemiological, information from the UN and WHO, and from other countries may all combine to provide a basis for comparison.

The work of Dr Pérez-Gómez refers predominantly to illicit drugs but sometimes also to alcohol, for which there are either very little or no official data, only estimates and a little survey data. Some of this information relates to policing and law enforcement but, as with other crime data collection, there are difficulties in determining the attributable fractions.

Estimates

It is often necessary to rely on estimates and expert opinion, which are the least authoritative level of evidence, and very far from the gold standard. Policy implications from these sources require cautious consideration.

As base-line cost data are currently being collected by CICAD in six countries, it should be possible to undertake avoidable cost estimates in some of those countries. When the initial base-line costs become available, they should be examined to determine whether the avoidable costs methodology could be applied to some of the identified cost categories.

Evaluation of prevention programs

These programs are reportedly usually very small and not very effective. Budgets for such programs are usually not identified, and they appear to be conducted independently of state and national policies. This makes the generalisability of the outcomes of such programs very difficult, in terms of their utility in determining avoidable costs.

Researchers

To undertake avoidable costs studies, it will be of great advantage to use the experience of researchers who have become familiar with the collection and use of cost data in the study countries.

The following tables from the workshop presentation of Dr Pérez-Gómez summarise CICAD's views of the relevance of illicit drugs and alcohol to the aggregate and avoidable social costs of substance abuse in Central and South America.

Table 23 – Relevance of avoidable costs to the health and welfare system in Central and South American countries

	Alcohol			Illicit drugs		
	High	**Fair**	**Low**	**High**	**Fair**	**Low**
Medical	Costa Rica, Argentina, Uruguay	Mexico	Colombia, Peru, Venezuela	Argentina	Costa Rica, Uruguay	Most
Hospital	El Salvador, Mexico	Costa Rica, Uruguay, Argentina	Most	El Salvador	Mexico, Uruguay, Costa Rica	Most
Nursing homes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Pharmaceuticals			All		Argentina, Brazil	Most
Ambulances			All			All
Research			All		Costa Rica, Peru, Brazil, Mexico	Most
Prevention			All			All
Social reinsertion			All		Peru, Colombia, Mexico, Costa Rica	Central American countries, Caribbean countries
Welfare administration			All			All
Treatment		Most	Ecuador, Panama, Caribbean countries	El Salvador	Most	Ecuador, Panama, Caribbean countries

Note: n.a. denotes not available

Table 24 – Attributed relevance of alcohol and drugs to the commission of specific crimes/offences in Central and South American countries

	Illicit drugs only			Alcohol only			Alcohol plus illicit drugs		
	High	**Fair**	**Low**	**High**	**Fair**	**Low**	**High**	**Fair**	**Low**
Violent	Colombia, Central American countries, Peru, Brazil	Mexico	Caribbean countries	All			All		
Property		All				All		All	
Drug offences	Barbados Caribbean countries	El Salvador, Colombia	Most	n.a.	n.a.	n.a.	Barbados, Caribbean countries	Most	
Breaches	Barbados, Caribbean countries	Most		All			All		
Disorder	Colombia, Peru, Mexico, El Salvador								
Drink driving				All			Colombia	Most	Caribbean countries
Other	Colombia, Brazil		Most						

Note: n.a. denotes not available

Table 25 – Relevance of illicit drugs to the judiciary system in Central and South American countries

	High	Fair	Low
Law enforcement	Colombia, Peru, Mexico, Venezuela	Bolivia, Ecuador, Panama, Brazil	Belize, Guyanas
Criminal courts	Colombia, Peru, Brazil	Caribbean countries, Central American countries, Venezuela	Belize, Guyanas
Prisons	Colombia, Peru, Central American countries	Argentina, Chile, Caribbean countries	Uruguay, Panama, Bolivia
Customs	Colombia, Peru, Brazil, Ecuador, Mexico	Panama, Haiti, Dominican Republic, Venezuela, Central American countries, Caribbean countries	Argentina, Panama, Bolivia
Organised crime	Colombia, Peru, Brazil, Mexico, El Salvador	Central American countries, Jamaica	Argentina, Caribbean countries, Chile, Bolivia
Forgone product			All
Property theft		Argentina, Chile, Brazil	Most
Violence	Colombia, Peru, Caribbean countries, Brazil	Jamaica, Venezuela	Argentina, Uruguay, Ecuador, Bolivia
Money laundering	Colombia, Peru, Caribbean countries	Brazil, Venezuela	Argentina, Colombia, Uruguay, Ecuador