Drug harms in the UK: a multicriteria decision analysis

David J Nutt, Leslie A King, Lawrence D Phillips, on behalf of the Independent Scientific Committee on Drugs

Summary

Background Proper assessment of the harms caused by the misuse of drugs can inform policy makers in health, policing, and social care. We aimed to apply multicriteria decision analysis (MCDA) modelling to a range of drug harms in the UK.

Method Members of the Independent Scientific Committee on Drugs, including two invited specialists, met in a 1-day interactive workshop to score 20 drugs on 16 criteria: nine related to the harms that a drug produces in the individual and seven to the harms to others. Drugs were scored out of 100 points, and the criteria were weighted to indicate their relative importance.

Findings MCDA modelling showed that heroin, crack cocaine, and metamfetamine were the most harmful drugs to individuals (part scores 34, 37, and 32, respectively), whereas alcohol, heroin, and crack cocaine were the most harmful to others (46, 21, and 17, respectively). Overall, alcohol was the most harmful drug (overall harm score 72), with heroin (55) and crack cocaine (54) in second and third places.

Interpretation These findings lend support to previous work assessing drug harms, and show how the improved scoring and weighting approach of MCDA increases the differentiation between the most and least harmful drugs. However, the findings correlate poorly with present UK drug classification, which is not based simply on considerations of harm.

Introduction

Drugs including alcohol and tobacco products are a major cause of harms to individuals and society. For this reason, some drugs are scheduled under the United Nations 1961 Single Convention on Narcotic Drugs and the 1971 Convention on Psychotropic Substances. These controls are represented in UK domestic legislation by the 1971 Misuse of Drugs Act (as amended). Other drugs, notably alcohol and tobacco, are regulated by taxation, sales, and restrictions on the age of purchase. Newly available drugs such as mephedrone (4-methylmethcathinone) have recently been made illegal in the UK on the basis of concerns about their harms, and the law on other drugs, particularly cannabis, has been toughened because of similar concerns.

To provide better guidance to policy makers in health, policing, and social care, the harms that drugs cause need to be properly assessed. This task is not easy because of the wide range of ways in which drugs can cause harm. An attempt to do this assessment engaged experts to score each drug according to nine criteria of harm, ranging from the intrinsic harms of the drugs to social and health-care costs. This analysis provoked major interest and public debate, although it raised concerns about the choice of the nine criteria and the absence of any differential weighting of them.

To rectify these drawbacks we undertook a review of drug harms with the multicriteria decision analysis (MCDA) approach. This technology has been used successfully to lend support to decision makers facing complex issues characterised by many, conflicting objectives—eg, appraisal of policies for disposal of nuclear waste. In June, 2010, we developed the multicriteria model during a decision conference, which is a facilitated workshop attended by key players, experts, and specialists who work together to create the model and provide the data and judgment inputs.

Methods

Study design

The analysis was undertaken in a two-stage process. The choice of harm criteria was made during a special meeting in 2009 of the UK Advisory Council on the Misuse of Drugs (ACMD), which was convened for this purpose. At this meeting, from first principles and with the MCDA approach, members identified 16 harm criteria (figure 1). Nine relate to the harms that a drug produces in the individual and seven to the harms to others both in the UK and overseas. These harms are clustered into five subgroups representing physical, psychological, and social harms. The extent of individual harm is shown by the criteria listed as to users, whereas most criteria listed as to others take account indirectly of the numbers of users. An ACMD report explains the process of developing this model.

In June, 2010, a meeting under the auspices of the Independent Scientific Committee on Drugs (ISCD)—a new organisation of drug experts independent of government interference—was convened to develop the MCDA model and assess scores for 20 representative drugs that are relevant to the UK and which span the range of potential harms and extent of use. The expert group was formed from the ISCD expert committee plus two external experts with specialist knowledge of...
legal highs (webappendix). Their experience was extensive, spanning both personal and social aspects of drug harm, and many had substantial research expertise in addiction. All provided independent advice and no conflicts of interest were declared. The meeting’s facilitator was an independent specialist in decision analysis modelling. He applied methods and techniques that enable groups to work effectively as a team, enhancing their capability to perform, thereby improving the accuracy of individual judgments. The group scored each drug on each harm criterion in an open discussion and then assessed the relative importance of the criteria within each cluster and across clusters. They also reviewed the criteria and the definitions developed by the ACMD. This method resulted in a common unit of harm across all the criteria, from which a new analysis of relative drugs harms was achieved. Very slight revisions of the definitions were adopted, and panel 1 shows the final version.

**Scoring of the drugs on the criteria**

Drugs were scored with points out of 100, with 100 assigned to the most harmful drug on a specific criterion. Zero indicated no harm. Weighting subsequently compares the drugs that scored 100 across all the criteria, thereby expressing the judgment that some drugs scoring 100 are more harmful than others.

In scaling of the drugs, care is needed to ensure that each successive point on the scale represents equal increments of harm. Thus, if a drug is scored at 50, then it should be half as harmful as the drug that scored 100. Because zero represents no harm, this scale can be regarded as a ratio scale, which helps with interpretation of weighted averages of several scales. The group scored the drugs on all the criteria during the decision conference.

Consistency checking is an essential part of proper scoring, since it helps to minimise bias in the scores and encourages realism in scoring. Even more important is the discussion of the group, since scores are often changed from those originally suggested as participants share their different experiences and revise their views. Both during scoring and after all drugs have been scored on a criterion, it is important to look at the relativities of the scores to see whether there are any obvious discrepancies.

**Weighting of the criteria**

Some criteria are more important expressions of harm than are others. More precision is needed, within the context of MCDA, to enable the assessment of weights on the criteria. To ensure that assessed weights are meaningful, the concept of swing weighting is applied. The purpose of weighting in MCDA is to ensure that the units of harm on the different preference scales are equivalent, thus enabling weighted scores to be compared and combined across the criteria. Weights are, essentially, scale factors.

MCDA distinguishes between facts and value judgments about the facts. On the one hand, harm expresses a level of damage. Value, on the other hand, indicates how much that level of damage matters in a particular context. Because context can affect assessments of value, one set of criterion weights for a particular context might not be satisfactory for decision making in another context. It follows then, that two stages have to be considered. First, the added harm going from no harm to the level of harm represented by a score of 100 should be considered—ie, a straightforward assessment of a difference in harm. The next step is to think about how much that difference in harm matters in a specific context. The question posed to the group in comparing the swing in harm from 0 to 100 on one scale with the swing from 0 to 100 on another scale was: “How big is the difference in harm and how much do you care about that difference?”

During the decision conference participants assessed weights within each cluster of criteria. The criterion within a cluster judged to be associated with the largest swing weight was assigned an arbitrary score of 100. Then, each swing on the remaining criteria in the cluster was judged by the group compared with the 100 score, in terms of a ratio. For example, in the cluster of four criteria under the category physical harm to users, the swing weight for drug-related mortality was judged to be the largest difference of the four, so it was given a weight of 100. The group judged the next largest swing in harm to be in drug-specific
mortality, which was 80% as great as for drug-related mortality, so it was given a weight of 80. Thus, the computer multiplied the scores for all the drugs on the drug-related mortality scale by 0.8, with the result that the weighted harm of heroin on this scale became 80 as compared with heroin’s score of 100 on drug-specific mortality. Next, the 100-weighted swings in each cluster were compared with each other, with the most harmful drug on the most harmful criterion to users compared with the most harmful drug on the most harmful criterion to others. The result of assessing these weights was that the units of harm on all scales were equated. A final normalisation preserved the ratios of all weights, but ensured that the weights on the criteria summed to 1.0. The weighting process enabled harm scores to be combined within any grouping simply by adding their weighted scores. Dodgson and colleagues provide further guidance on swing weighting. Scores and weights were input to the Hiview computer program, which calculated the weighted scores, provided displays of the results, and enabled sensitivity analyses to be done.

Role of the funding source
The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study, and had final responsibility for the decision to submit for publication.

Results
Figure 1 shows the 16 identified harm criteria. Figure 2 shows the total harm score for all the drugs and the part-score contributions to the total from the subgroups of harms to users and harms to others. The most harmful drugs to users were heroin (part score 34), crack cocaine...
and metamfetamine (32), whereas the most harmful to others were alcohol (46), crack cocaine (17), and heroin (21). When the two part-scores were combined, alcohol was the most harmful drug followed by heroin and crack cocaine (figure 2).

Another instructive display is to look at the results separately for harm to users and to others, but in a two-dimensional graph so that the relative contribution to these two types of harm can be seen clearly (figure 3). The most harmful drug to others was alcohol by a wide margin, whereas the most harmful drug to users was crack cocaine followed closely by heroin. Metamfetamine was next most harmful to users, but it was of little comparative harm to others. All the remaining drugs were less harmful either to users or to others, or both, than were alcohol, heroin, and crack cocaine (figure 3). Because this display shows the two axes before weighting, a score on one cannot be compared with a score on the other, without knowing their relative scale constants.

Figure 4 shows the contributions that the part scores make on each criterion to the total score of each drug. Alcohol, with an overall score of 72, was judged to be most harmful, followed by heroin at 55, then crack cocaine with a score of 54. Only eight drugs scored, overall, 20 points or more. Drug-specific mortality was a substantial contributor to five of the drugs (alcohol, heroin, γ hydroxybutyric acid [GHB], methadone, and butane), whereas economic cost contributed heavily to alcohol, heroin, tobacco, and cannabis.

Discussion
The results from this MCDA analysis show the harms of a range of drugs in the UK. Our findings lend support to the conclusions of the earlier nine-criteria analysis undertaken by UK experts and the output of the Dutch addiction medicine expert group. The Pearson correlation coefficient between Nutt and colleagues’ 2007 study and the new analysis presented here for the 15 drugs common to both studies is 0·70. One reason for a less-than-perfect correlation is that the scores from Nutt and colleagues’ previous study were based on four-point ratings (0=no risk, 1=some risk, 2=moderate risk, and 3=extreme risk). The ISCD scoring process was based on 0–100 ratio scales, so they contain more information than the ratings do.

Throughout Nutt and colleagues’ 2007 paper, harm and risk are used interchangeably, but in the ISCD work, risk was not considered because it is susceptible to varying interpretations. For example, the British Medical Association defines risk as the probability that something unpleasant will happen. Thus, assessors from Nutt and colleagues’ 2007 work might have interpreted their rating task differently from the scoring task of the ISCD experts. Furthermore, in Nutt and co-workers’ 2007 study, ratings were simply averaged across the nine criteria (called parameters in the report), three each for physical harm, dependence, and social harms, whereas differential weights were applied to the criteria in this ISCD study, as is shown in the key of

Figure 2: Drugs ordered by their overall harm scores, showing the separate contributions to the overall scores of harms to users and harm to others.

The weights after normalisation (0–100) are shown in the key (cumulative in the sense of the sum of all the normalised weights for all the criteria to users, 46; and for all the criteria to others, 54). CW=cumulative weight. GHB=γ hydroxybutyric acid. LSD=lysergic acid diethylamide.
Despite these many differences between the two studies, there is some degree of linear association between both sets of data.

The correlations between the Dutch addiction medicine expert group and ISCD results are higher: 0.80 for individual total scores and 0.84 for population total scores. As with Nutt and colleagues' 2007 study, the Dutch experts applied four-point rating scales to 19 drugs. However, they used five criteria: acute toxicity, chronic toxicity, addictive potency, social harm at individual level, and social harm at population level. Simple averages produced two overall mean harm ratings, one each for individuals and for populations. The probable explanation for the greater correlation between the ISCD and Dutch data lies in the greater relative ranges of the overall results than in Nutt and co-workers' 2007 study. The highest and lowest overall harm scores in the ISCD study are 72 for alcohol and 5 for mushrooms, which is a ratio of about 14:1; whereas in Nutt and colleagues' study it was a ratio of just over 3:1, from 2.5 for heroin to 0.8 for khat. The highest and lowest scores for the Dutch individual ratings were 2.63 for crack cocaine and 0.40 for mushrooms, which is a ratio of 6.6:1; and for the population ratings 2.41 for crack cocaine and 0.31 for mushrooms, which is a ratio of 7.8:1. The ratio scaling in the ISCD study spanned a wider range, making the three most harmful drugs—alcohol, heroin, and crack cocaine—much more harmful relative to the other drugs than can be expressed with rating scales, so that additional information stretched the scatterplot in one dimension, making it seem more linear. Additionally, because the Dutch scale attributes only a quarter of the scores to social factors, whereas in the ISCD scoring these factors comprise nearly half of the scores (seven of 16 criteria), drugs such as alcohol which have a major effect will rank more highly in the ISCD analysis, with tobacco ranked lower because its harms are mainly personal.

The correlations between the ISCD overall scores and the present classification of drugs based on revisions to the UK Misuse of Drugs Act (1971) is 0.04, showing that there is effectively no relation. The ISCD scores lend support to the widely accepted view that alcohol is an extremely harmful drug, both to users and society; it scored fourth on harms to users and top for harms to society, making it the most harmful drug overall. Even in terms of toxic effects alone, Gable has shown that, on the basis of a safety ratio, alcohol is more lethal than many other drugs.
illicit drugs, such as cannabis, lysergic acid diethylamide (LSD), and mushrooms.

The MCDA process provides a powerful means to deal with complex issues that drug misuse presents. The expert panel’s scores within one criterion can be to some extent validated by reference to published work. For example, we compared the 12 substances in common between this study and those in Gable’s study, who for 20 substances identified a safety ratio—the ratio of an acute lethal dose to the dose commonly used for non-medical purposes. The log10 of that ratio shows a correlation of 0.66 with the ISCD scores on the criterion drug-specific mortality, providing some evidence of validity of the ISCD input scores.

We also investigated drug-specific mortality estimates in studies of human beings. These estimates show a strong correlation with the group input scores: the mean fatality statistics from 2003 to 2007 for five substances (heroin, cocaine, amphetamines, MDMA/ecstasy, and cannabis) show correlations with the ISCD lethality scores of 0.98 and 0.99, for which the substances recorded on the death certificates were among other mentions or sole mentions, respectively.

A comparison of the ICSD experts’ ratings on the dependence criterion with lifetime dependence reported in the US survey by Anthony and co-workers showed a correlation of 0.95 for the five drugs—tobacco, alcohol, cannabis, cocaine, and heroin—that were investigated in both studies, showing the validity of the MCDA input scores for those substances.

Drug-specific and drug-related harms for some drugs can be estimated from health data and other data that show alcohol, heroin, and crack cocaine as having much larger effects than other drugs. Social harms are harder to ascertain, although estimates based on road traffic and other accidents at home, drug-related violence, and costs to economies in provider countries (eg, Colombia, Afghanistan, and Mexico) have been estimated. Police
Panel 2: Research in context

Systematic review
We analysed the data obtained from a multicriteria decision analysis (MCDA) conference on drug harms. The harms were assessed according to a new set of 16 criteria developed by the Advisory Council on the Misuse of Drugs (the UK Government committee on drug misuse). A panel of drug-harm experts was convened to establish scores for 20 representative drugs that are relevant to the UK and which span the range of potential harms and extent of use. Participants scored the relative harms of each drug on each of 16 criteria, and then assessed criterion weights to ensure that units of harm were equivalent across all criteria. Calculation of weighted scores provided a composite score on two dimensions, harm to the individual and harm to society, and an overall weighted harm score.

Interpretation
These findings lend support to earlier work from both UK and Dutch expert committees on assessment of drug harms, and show how the improved scoring and weighting approach of MCDA increases the differentiation between the most and least harmful drugs. On the basis of these data it is clear that the present UK drug classification system is not simply based on considerations of harm.

records lend support to the effect of drug dealing on communities and of alcohol-related crime.19 However, data are not available for many of the criteria, so the expert group approach is the best we can provide. The many high correlations (of our overall results with those of the Dutch addiction medicine expert group, and of some of our input scores with objective data) provide some evidence of the validity of our results.

The issue of the weightings is crucial since they affect the overall scores. The weighting process is necessarily based on judgment, so it is best done by a group of experts working to consensus. Although the assessed weights can be made public, they cannot be cross-validated with objective data. However, the effect of varying the weightings can be explored in the computer program through sensitivity analysis. For example, we noted that it would be necessary to increase the weight on drug-specific mortality or on drug-related mortality by more than 15 of 100 points before heroin displaced alcohol in first position of overall harm. A similarly large change in the weight on drug-specific damage would be needed, from about 4% to slightly more than 70%, for tobacco to displace alcohol at first position. And an increase in the weight on harm to users from 46% to nearly 70% would be necessary for crack cocaine to achieve the overall most harmful position. Extensive sensitivity analyses on the weights showed that this model is very stable; large changes, or combinations of modest changes, are needed to drive substantial shifts in the overall rankings of the drugs. Future work will explore these weightings with use of other groups—both expert panels and those from the general public.

Limitations of this approach include the fact that we scored only harms. All drugs have some benefits to the user, at least initially, otherwise they would not be used, but this effect might attenuate over time with tolerance and withdrawal. Some drugs such as alcohol and tobacco have commercial benefits to society in terms of providing work and tax, which to some extent offset the harms and, although less easy to measure, is also true of production and dealing in illegal drugs.19 Many of the harms of drugs are affected by their availability and legal status, which varies across countries, so our results are not necessarily applicable to countries with very different legal and cultural attitudes to drugs. Ideally, a model needs to distinguish between the harms resulting directly from drug use and those resulting from the control system for that drug. Furthermore, they do not relate to drugs when used for prescription purposes. Other issues to explore further include building into the model an assessment of polydrug use, and the effect of different routes of ingestion, patterns of use, and context.19 Finally, we should note that a low score in our assessment does not mean the drug is not harmful, since all drugs can be harmful under specific circumstances.

In conclusion, we have used MCDA to analyse the harms of a range of drugs in relation to the UK (panel 2). Our findings lend support to previous work in the UK and the Netherlands, confirming that the present drug classification systems have little relation to the evidence of harm. They also accord with the conclusions of previous expert reports11,18 that aggressively targeting alcohol harms is a valid and necessary public health strategy.

Contributors
DJN designed and participated in the study. LAK participated in the study. LDP participated in the running of the study and analysed data. CJW, LDP, and DJN wrote the report and responded to referees’ comments.

Conflicts of interest
DJN and LAK received travel expenses to attend the decision conference meeting. LAK is a consultant to the Department of Health and the EMCDDA. LDP is a director of Facilitations Limited, which paid him a consulting fee because it was the company engaged by the Centre for Crime and Justice Studies to run the study and analyse the data.

Acknowledgments
This study is funded by the Centre for Crime and Justice Studies (UK). Yuji Wu assisted with some of the data analyses.

References


