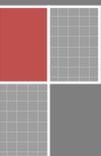


# **A review of estimation methods on prevalence of drug abuse population in Hong Kong**

## ***Executive Summary***

Prepared by  
***Policy 21 Limited***  
February 2013



1. There is no universally accepted method to accurately measure the size of the drug abuse population in a country or territory. Traditional estimation methods such as territory-wide surveys tend to be inaccurate due to large sampling and possibly under-reporting errors.

2. As such, Hong Kong all along adopts a multi-dimensional approach to monitor the local drug abuse situation including drug statistics obtained in the Central Registry of Drug Abuse (CRDA), research studies and student survey. While such approach has proved to be efficient in monitoring the trends of drug abuse, which is crucial for policy formulation, it is unable to produce an accurate measure of the prevalence of drug abuse.

3. Furthermore, there are diverse views on the definition of drug abusers worldwide. As pointed out by researchers, how drug abusers should be defined was closely linked to the purpose of compiling the estimate.

4. The objective of the present study is to recommend a more suitable definition of drug abusers in the context of Hong Kong with reference to international criteria; and to recommend estimation method(s) for estimating the overall and sub-group drug abuse population in Hong Kong.

## **Definition**

5. The intended use of statistics on drug abusers is closely related to how drug abusers should be defined. If the purpose is to assess possible treatment needs, the definition should be related to potential clients of treatment agencies, either now or in the future. If it is to assess demand, and thus the illicit market, for a drug such as heroin, all heroin users should be included, regardless of whether they are likely to seek treatment. Furthermore, if the dominant aim is to suppress all drug use, a definition of drug abusers that includes all types of drug use would be more appropriate. If the emphasis is placed on “harm reduction”, the number of drug users should include those who are clients or potential clients for harm reduction intervention programmes.

### ***Practices in Hong Kong and other countries***

6. In Hong Kong, for the purpose of reporting to CRDA, “*a drug abuser is defined to be a person who has come into contact with a particular agency and is known or suspected to have taken substances during the four weeks before the date of contact, and the substances if taken harm or threaten to harm the physical, mental or social well-being of an individual, in doses above or for periods beyond those normally regarded as therapeutic.*”

7. In countries like Canada, drugs are considered in the broadest sense to include any substance which is taken to change the way the body and/or mind function, and its continued use is referred to as drug abuse. In the United Kingdom, the definition of drug abuse covers not only illegal drugs but also other drugs the misuse of which is harmful to health. The frequency and length of time in which drugs are used are more relaxed. In USA, drugs cover only the most popular and commonly used drugs and drug abusers normally refer to non-institutional population aged 12 or over taking drugs. It is worth noting that non-medical use of prescription drugs is included as one of the conditions for the classification of drug abuse.

### ***Relevant factors to consider***

8. Recognizing youth drug abuse is a complex problem, the Government has adopted a holistic approach in its anti-drug policy, following a five pronged approach comprising preventive education and publicity, treatment and rehabilitation, legislation and law enforcement, external cooperation and research.

9. It is thus desirable to adopt a definition that covers different categories of drug abusers such that information obtained is useful for planning, policy formulation, programme evaluation and research purposes. Furthermore, due considerations have to be given to the feasibility of “operationalizing” the definitions in data collection and programme implementation, such that respondents in surveys conducted to estimate the prevalence rate of drug abuse are able to answer questions asked, as well as frontline workers responsible are able to report cases to relevant authorities and/or CRDA. Needless to say, such definitions have to be compatible with existing rules and regulations adopted in law enforcement.

### ***Reporting agencies***

10. CRDA data are derived from “institutionally visible” population when they are contacted by the legal, medical or social systems. Any definition adopted should take into account the interests, norms and values of these three systems.

11. There are a number of reporting agencies to the CRDA. Each reporting agency provides different services to different types of drug abusers, following procedures and service criteria applicable to the agency concerned. As recommended in the Report on Review of Central Registry of Drug Abuse, the definitions adopted should be simple enough to be used by all reporting agencies.

## **Proposed changes to the definition of drug abusers**

### ***Overview***

12. Several parameters are relevant in defining drug abusers over a certain period of time, which include the specific substance used, legal status of drugs or substances used, administration method, frequency of use and clinical diagnosis. To evaluate the impact of anti-drug programmes and to plan for the provision of treatment and preventive programmes, it is desirable to adopt a definition that covers different categories of drug abusers.

### ***Specific substances used***

13. *It is recommended that the definition of drugs and substances adopted by CRDA should continue to be adopted.* Given the broad coverage of drugs and substances, some frontline staff may inevitably have difficulties in deciding whether or not the use of certain drugs or substances should be reported or not. To facilitate reporting agencies in reporting data to the CRDA, a list of commonly abused drugs and substances are given to reporting agencies. *As drug abuse patterns change quite often, in response to the availability of new, synthetic drugs, it is recommended that the list should be updated more frequently. In particular, reporting agencies should be encouraged to report the use of drugs or substances that are considered harmful to the physical, mental or social well-being of an individual, including those not shown in the list of commonly abused drugs and substances, especially the nonmedical use of prescription drugs.*

### *Frequency of use*

14. At present, a drug abuser is defined as a person who has come into contact with a particular agency and is known or suspected to have taken substances during the four weeks before the date of contact. The CRDA review in 2001 considered the four-week guideline reasonable to ensure that memories of drug abusers were still fresh to provide accurate details of their drug abuse experience. However, some drug abusers contacted by reporting agencies would only abuse drug in certain times of the year like school holidays and festive seasons. If the reference period of four weeks is strictly applied, the CRDA may not be able to capture drug abusers contacted by reporting agencies in other times of the year (e.g. school days).

15. Indeed, the criteria adopted in the Diagnostic and Statistical Manual of Mental Disorders (DSM) for dependence on or abuse of specific psychotropic substance groups use a reference period of twelve months. In Canada, continued use of drug is referred to as drug abuse when it becomes a problem or when it results in negative consequences for the person using it, with no restriction as to the period of drug use. Nevertheless, in collecting data through telephone interviews on reported abuse of drugs, the reference period of past twelve months prior to enumeration is used. Similar practices are also adopted in countries like UK and USA in collecting drug abuse statistics through surveys of the population.

16. Overall speaking, the current reference period of four weeks should be maintained to ensure comparability of data and more accurate reporting by drug abusers. Indeed, according to the 2001 review, the four-week guideline is not meant to be rigid but “only a reference time frame to ensure data compatibility and the accuracy of the information provided by the drug abusers”. For drug abusers who are found to be taking drugs irregularly but persistently, reporting agencies are requested to report such cases to CRDA as well.

17. At the same time, there is a case to explore further how to better capture those who have not, for various reasons, taken drugs during the four-week reference period. One possible direction is to consider adding a new reference period of one year if it is deemed feasible and useful after a thorough assessment. It is noteworthy that CRDA data is collected from 72 reporting agencies of various natures, including welfare organisations, law enforcement departments, drug treatment and rehabilitation centres, educational and healthcare institutions. While some agencies such as drug treatment and rehabilitation centres would ask their clients anyway about their drug abuse pattern beyond the four-week reference period in order to provide appropriate services, some others may find such questions challenging in view of the circumstances of engagement and the attitudes of respondents. More fundamentally, CRDA is intended to shed light on the drug trend across time rather than the exact size of drug abuse population. Unless the limitations of the estimate methods considered by this study can be satisfactorily resolved, a new reference period may not add much value to the current system.

18. To sum up, *while the current reference period of four weeks prior to contact by reporting agencies should be maintained, consideration can be given to introducing an additional reference period of one year.* Before introducing such a change, a thorough assessment of the feasibility and usefulness should be conducted. There should also be consultation with reporting agencies and users of the CRDA data.

### ***Legal status of the substance used***

19. No specific requirement is included in the current definition on the legal status of drugs or substances used. This is consistent with practices adopted in countries reviewed and the anti-drug policy of Government. Besides, for reporting agencies in the medical and social systems, their concern is on the service needs of their drug abuse clients, rather than the legal status of the drugs or substances their clients have taken. *It is suggested not to include the legal status of the substance used in the definition of drug abuser.*

### ***Clinical diagnosis***

20. In providing services to drug abusers, reporting agencies in the medical system would undoubtedly take steps to assess the clinical conditions of drug abusers, including whether they have drug dependence, physically or mentally. It is conceivable that reporting agencies in the social system would undertake similar assessment. For reporting agencies in law enforcement departments, they should have similar medical assessment of the health conditions of the drug abusers, such that special arrangements could be made when the drug abusers are put under their custodial care.

21. It is reckoned that gathering information on drug dependence and abuse will pose additional reporting burden on reporting agencies, at the risk of reducing incentives of reporting agencies to submit drug abuse data to the CRDA. There will be a need to train frontline staff responsible for reporting data to the CRDA on assessing drug dependence and abuse of their clients. The costs involved on the part of reporting agencies will not be insignificant. Thus, the expected benefits to be derived from having information on drug dependence and abuse should be weighed against the likely costs involved. *It is recommended that the benefits and costs of collecting information on drug dependence and abuse, following ICD codes and DSM diagnostic criteria, by reporting agencies should be examined and the feasibility investigated.*

22. It was noted by researchers that measures of drug dependence devised based on ICD codes and DSM diagnostic criteria were not commonly used by data sources in the UK. The applicability of such measures in Hong Kong should be thoroughly examined and if found suitable, promoted to all relevant data sources.

## **Estimation methods**

23. There are various methods available for estimating the prevalence of drug use, including population-based surveys, case-finding studies, capture-recapture methods, multiplier techniques, nomination techniques (including snowball sampling), synthetic estimates based on social or demographic variables assumed to have correlation with drug prevalence and a variety of more sophisticated statistical modeling approaches.

### ***Case-finding method***

24. Case-finding is a standard epidemiological method for obtaining an adequate number of cases for observation and research. As there is no single way or information source which can find all or at least most drug users, a combination of different strategies is needed. Multi-sources information is one method widely used to overcome the lack of completeness and representativeness of a single source and can provide estimates of the prevalence of drug users. When different sources are combined, there is considerable risk of over-estimating the total number of cases unless personal identifiers are available on a reliable and standardized way in all sources to enable the identification and removal of duplicates.

25. Case-finding is applicable to study drug use for several reasons. Firstly, drug use is rare. Secondly, as an illegal activity, it is largely hidden. Therefore, general population sample survey techniques will be too costly, inefficient, and may be ineffective for identifying drug users. Thirdly, a readily-made sampling frame or register does not exist, which, in part, is the reason for carrying out a case-finding study. The CRDA is indeed an example of the case-finding method.

### ***Capture-recapture method (CRM)***

26. The capture-recapture method involves “capturing” a random sample which is then “marked” and returned to its habitat. Subsequently, a second random sample is “recaptured” and the number of those marked from the first sample is observed. The ratio of marked sample size to the recaptured sample size is assumed to be the same as the ratio of the first captured sample to the total population.

27. It should be noted that the following assumptions of the method are important for its success:

- a) The population under study must be closed, in the sense that individuals do not enter or leave the population during the study period;
- b) The samples must be randomly selected and the probability of each individual being selected must be the same in each sample;
- c) The samples must be mutually independent.

28. As not all of the above conditions can be met in reality, researchers have attempted to use multiple data sources for circumstances where the probabilities of selection are not the same (condition b) above) or that the samples are not mutually independent (condition c) above).

#### ***Multiplier method (MM)***

29. The multiplier method involves applying a “multiplier” to a “benchmark” (the total of a sub-group of the drug-using population). The most commonly used “benchmark” is the total number of drug-related deaths (or mortality data). However, it is usually difficult for mortality data to show all deaths that are related to drug abuse. As mortality and hospitalization due to health problems arising from the use of drugs are not recorded as drug related incidents, this method is considered not applicable in Hong Kong.

#### ***Nomination method***

30. The principles involved in using the nomination method specifically to estimate the prevalence of drug abuse are the same as the multiplier method. The sampled members are asked to name or nominate drug-abusing acquaintances and to say whether these acquaintances have been in touch with drug treatment centres, health services or any other similar body, within a stipulated time period.

31. Due to confidentiality considerations, the nomination method which relies on a sample of drug abusers to nominate drug-abusing acquaintances, is not likely to be feasible. Reporting agencies are also reluctant to approach their clients to seek their consent for the release of personal information for the purposes of the study, for fear that doing so would ruin the relationship with their clients and increase reluctance on the part of their clients to report their information to them.

## **Simulation runs of the “capture re-capture method” (CRM)**

### ***Data source***

32. To test the CRM, the anonymous list of CRDA records is divided into different datasets. Based on the datasets, a number of CRM models were run.

- a) Stratification of the list into male and female;
- b) Twelve datasets of previously reported and newly reported cases in each of the twelve-year period from 2000 to 2011 for males and another twelve sub-sets for females;
- c) For each dataset, information contained in the list includes reported year, year of birth, ethnicity, type of substances and frequency of use, as well as an indication on the categories of reporting agencies.

### ***Two-source capture-recapture models***

33. Taking the CRDA as a one combined, single data source, the capture-recapture estimation was conducted.

### **Estimation by gender –**

34. One of the conditions for the capture-recapture approach is that the samples must be randomly selected and the probability of each individual being selected must be the same in each sample. However, it is believed that the probability of recapturing male and female drug abusers is likely to be different. To reduce heterogeneity, separate capture-recapture models were run for male and female drug abusers.

35. The precision level of the estimates obtained from the two-source capture-recapture models is quite high. The estimate ranges within  $\pm 1.8\%$  (for males) and  $\pm 5.3\%$  (for females) of the estimate at 95% confidence.

### **Estimation by gender and age group –**

36. To further minimize heterogeneity due to different recapture probabilities, separate capture-recapture models were run for drug abusers of different age-sex groups. Findings of the models would not only produce estimates of the number of drug abusers by age groups but also as a counter-check to the estimates computed above.

37. The precision level of the estimates obtained from the two-source capture-recapture models, with separate estimates computed for different age groups, is quite high. At 95% confidence, the estimate ranges within  $\pm 1.7\%$  (for males) and  $\pm 5.0\%$  (for females). This level of precision is considered quite acceptable.

**Estimation by gender and types of drugs –**

38. Drug abusers were classified into opiate and non-opiate users in the model. The precision level of the estimates obtained from the two-source capture-recapture models, with separate estimates computed for opiate and non-opiate drug abusers, is also quite high. At 95% confidence, the estimate ranges within  $\pm 1.5\%$  (for males) and  $\pm 4.1\%$  (for females).

39. Findings based on simulation runs described above show that the margin of estimation errors for a two-source CRM is very small.

***Four-source capture-recapture model***

40. Attempts have been made to minimize heterogeneity by estimating separately the number of drug abusers by gender, age, types of drugs used and ethnicity by classifying reporting agencies into four different categories. However, the precision level of the estimates obtained from the four-source capture-recapture models is quite low, and is considered not quite acceptable.

***Three-source capture-recapture model***

41. Attempt has made to reclassify reporting agencies such that those with similar recapture probabilities are grouped into one single source. Again, the precision level of the estimates obtained from the three-source capture-recapture models is quite low, and is considered not quite acceptable.

**Comparing estimates derived from different models**

42. Making use of an anonymous list of drug abusers extracted from the CRDA, different capture-recapture models were tested and compared. The salient findings are summarized below:

- a) Based on three-source and four-source capture-recapture models and using log linear regression techniques, separate estimates of the number of male and

female drug abusers were computed. However, no valid model could be selected for a number of years during the study period from 2001 to 2011. Besides, the margin of estimation errors is found to be unacceptably large;

- b) Based on two-source capture-recapture models, four sets of estimates were computed separately for male and female drug abusers for each year during the study period from 2001 to 2011. The margin of estimation errors is very small. *Thus, the two-source capture-recapture model is proposed to be adopted.*

43. By adopting the two-source capture-recapture models, different sets of estimates were obtained by computing four sets of separate estimates for gender only, gender-age, gender-ethnicity and gender-types of drugs used. For male drug abusers, the estimates derived by the capture-recapture models based on gender only and gender-ethnicity are almost the same. The estimates derived from capture-recapture model based on gender-age are slightly higher, and those derived from the capture-recapture model based on gender-types of drugs used are much higher. Nevertheless, the patterns of changes are quite similar for all four sets of estimates.

44. For female drug abusers, the estimates derived by the capture-recapture models based on gender and ethnicity are almost the same. The estimates derived from capture-recapture model based on gender-and-age are slightly higher during the period from 2002 to 2005 and are almost the same as those derived based on gender and ethnicity for the years 2006 – 2011. The estimates derived from the capture-recapture model based on drug types are higher except for the year of 2004. Nevertheless, the patterns of changes are quite similar for all four sets of estimates.

### **Recommended methodology for estimating the number of drug abusers**

45. As discussed above, estimates derived from the four-source and three-source capture-recapture models are subject to unacceptably large margins of estimation errors, while precision levels for those derived from the two-source capture-recapture models, with separate estimates by gender, gender-age, gender-ethnicity and gender-types of drugs used, are much higher. As noted by researchers, using many sources capture-recapture techniques and log-linear modeling was considered too complex. On the other hand, the use of two-source capture-recapture techniques when appropriate, allowed for more simplified analysis. *In view of the above findings, it is proposed to adopt the two-source capture-recapture methodology in estimating the number of drug abusers.*

46. Given that the three models produce estimates which are quite close to each other, *it is recommended that estimates based on the two-source capture-recapture with breakdowns by gender, ethnicity and drug type should continue to be compiled, to*

*provide a rough indication of the size of sub-groups of interest and relevant to policy formulation (e.g. the number of non-opiate drug abusers). As re-runs of the three models, based on updated data obtained from the CRDA may be performed in-house, the cost involved is not likely to be significant.*

47. Attempts may be made to re-run the CRM models with more detailed breakdown by say gender-age groups (e.g. male drug abusers aged 30 – 35). While such estimation is possible, the estimates derived will be subject to higher estimation errors due to smallness of the sample size and tend to be unstable. It is suggested that the overall estimates by three broad age groups (namely below 21, 21 – 50 and above 50) and separately by gender above should be used as control totals.

48. Three other methods in estimating the size of the drug abuse, namely the case-finding, multiplier and nomination method, have been examined in this research. As noted above, the CRDA is in fact a notable example of the case-finding method. For the multiplier method, it has been used to serve as a counter-check of the reliability of estimates derived from the capture-recapture method, making use of statistics on the number of arrests related to drug offenses. As discussed above, statistics on drug related arrests or mortality arising from the use of drugs suffer from a number of limitations rendering such statistics unsuitable for the purposes of multiplier method. Reluctance on the part of reporting agencies to seek consent from drug abusers to participate in surveys or interviews required for the purposes of the nomination method also severely reduce the usefulness of this method in estimating the size of the drug abuse population. *In the circumstance, the multiplier and nomination method are not recommended for the purposes of estimating the size of the drug abuse population.*

## **Limitations**

49. The success of the CRM depends on a number of assumptions which cannot be made in reality. For example, the drug abuse population in Hong Kong can hardly be closed, as there are individuals leaving through mortality or migration. Each year, there are people beginning to take drugs, thus joining the drug abuse population. Besides, the samples of drug abusers selected from the drug abuse population through the use of the CRDA are not random samples. Furthermore, the probability of reporting agencies contacting again the drug abusers who have been reported by, say, the Correctional Services Department will be much lower than that for drug abusers who have been reported by, say, outreach social workers, as drug abusers in the former case will mostly likely be in the custodial care provided by the Correctional Services Department and will unlikely be contacted again by other reporting agencies. In other words, the samples are not mutually independent.

50. To meet the condition that the population under study must be closed, researchers have attempted to shorten the reference period. In the present study, the reference period

adopted is one year, which hopefully would help minimize the mobility into or out of the drug abuse population. To minimize the differences in the probability of selection, separate estimates have been compiled for different gender groups, age-sex groups, ethnicity groups and types of drugs used. By modeling the CRM separately on, say, male abusers aged under 21, 21 to 50 and over 50, variations within each age-sex group in the probability of selection will be much smaller than variations for all male drug abusers or all drug abusers.

51. Attempts have also been made to minimize the impact of the fact that the samples are not mutually independent by separately modeling the CRM on drug abusers reported by different categories of reporting agencies, using the four-source and three-source CRM. The model findings, however, are found to be unstable and the estimates are subject to large estimation errors.

52. As it is not possible to derive an independent estimate of drug abusers by contacting the drug abusers, it is not possible to assess the impact of the deviations from the assumptions required of the CRM. Nevertheless, it may be worth noting that the estimates derived from the two-source CRM, for different gender groups, age-sex groups, ethnicity groups and types of drugs, are not far apart. In other words, the impact of the partial deviations from the assumption that the probability of selection should be the same is not large.

53. Nevertheless, readers should note the limitations spelt out above that the assumptions required of the CRM are in practice not met. The estimates presented in this study could at best give a rough estimate of the size of the drug abuse population. Thus, the estimates derived from the present study should serve as a reference only. The estimates should not be taken as the basis for resource planning. Given that the CRM is consistently applied, the estimates could also provide an indication of the changes over time. There is still a continued need to publicize the use of the CRDA and the confidentiality protection it has afforded to individuals affected and to urge reporting agencies to report fully drug abusers they have contacted, such that the extent of under-reporting of drug abusers to the CRDA would be reduced.