DATA
FOR
SOLUTIONS

Size Estimation for High Risk Groups
Moldova

Prepared for: Moldova UNAIDS Country Office
DRAFT: February 22, 2010
Donna F. Stroup, PhD, MSc

Moldova Size Estimation Mission

Background and Objectives of Mission

UNAIDS regularly reports the status of the HIV epidemic in large yearly reports (EpiUpdate, Global Report). These reports are global in nature, reflecting regional or country specific trends in the epidemic. For many countries in the world, the data required for these types of national and global reports require information about high risk populations. Historically, the size of high risk populations has been difficult to estimate due to issues of stigma, risk of harm to participants, and/or definition of risk behaviour.

The overall objective of the Mission in Moldova was to produce national estimations of sizes of most at-risk populations (MARPs) based on available data obtained through the NSU, the multiplier, and the unique object (capture-recapture) methods.

Specific Objective**:**

- ✓ Verify data obtained from the NSU and the estimations of hidden populations based on the co-efficient between estimated network components sizes and official statistical data on the respective known groups;
- ✓ Assist the national working group on interpreting the respect factor when analyzing the data obtained in the framework of the GPS through the NSU method;
- ✓ Compare the data on IDU obtained through the NSU method with the multiplier and capture recapture methods data;
- ✓ Assist the national technical work group in producing national estimations, ensuring transfer of expertise and strengthening of capacities in the process;
- ✓ Ensure technical backstopping to the national technical work group in producing estimations on sizes of SW and MSM.

Background on HIV and AIDS risk in Moldova

In the Republic of Moldova, 4996 HIV cases have been identified and confirmed throughout 1987-2008, including 1535 in Transnistria (1). In the last three years alone, 43 per cent of all cases have been reported. Based on UNAIDS estimations, 8865 persons (including 51 children) were living with HIV/AIDS at the end of 2007 in Moldova.

The onset of the epidemic in Moldova is considered to be 1996,[1] driven by Injecting Drug Users (IDUs) who spread the virus through sharing contaminated needles. In 2005, however, newly registered HIV cases associated with sexual transmission outnumbered those associated with injecting drug use: out of 795 new HIV cases reported in 2008, about 75,8% were sexually transmitted. Considering the high economic migration rates, this phenomenon could become a determinant in the future evolution of HIV in Moldova.[2] The shift in the structure of newly reported HIV cases according to their route of transmission increases the vulnerability of women. The share of HIV-infected women increased from 26,7 per cent in 2001 to 43,7 per cent in 2008.

According to the results of the HIV prevalence survey conducted in Most-at-risk Populations among beneficiaries of harm reduction programs in 2007, the HIV prevalence was 21% for IDUs, 11% for Commercial Sex Workers (CSWs), and 4.8% for Men having Sex with Men (MSM).[2] These data show an increase in HIV prevalence from 2004 among IDUs and MSM (Tables 1, 2, and 3).

The HIV epidemic in the Republic of Moldova is still considered to be concentrated mostly in IDUs population with sign of spreading into the general population. The high level of external migration for the left bank of the Dniester River creates favourable conditions for the rapid spread of HIV infection in the general population. The generalization of the epidemic is shown by data on annual rise in HIV incidence, and the geographic spread of the epidemic. Specifically, 41.64 per cent of all new cases diagnosed in 2008 were in rural areas versus 38.7 per cent reported in 2007.

Although the first study on risk behaviors in most-at-risk adolescents (MARA) was conducted in 2007-2008, there are no data estimates on the size of this group. The mandatory consent of parents/custodians when doing any medical procedure in underage groups is a major barrier in enrolling MARA in HIV prevalence studies conducted in the most-at-risk populations (MARP).

---

[1] National Drug Observatory, 2006
[2] Oxford Analytica 2008

Table 1 HIV prevalence in IDUs, Republic of Moldova, 2001- 2007

Data collection site	2001[3]		2003 - 2004[4]		2007[5]	
	Sample	HIV prev	Sample	HIV prev	Sample	HIV prev
Chisinau, capital city	**209**	**15.8 %**	**306**	**14.4 %**	**183**	**17.5 %**
Balti	184	60,3 %	230	36,5 %	145	44,8 %
Causeni	n/a	n/a	10	40,0 %	11	27,3 %
Donduseni	n/a	n/a	n/a	n/a	10	10,0 %
Edinet	n/a	n/a	7	14,3 %	20	15,0 %
Falesti	50	22,0 %	67	11,9 %	28	10,7 %
Orhei	13	23,1 %	44	2,3 %	21	0 %
Rezina	n/a	n/a	43	11,6 %	30	16,7 %
Soroca	87	1,15 %	116	0,0 %	41	0 %
Tiraspol	n/a	n/a	n/a	n/a	68	20,7%
Ungheni	n/a	n/a	47	2,3 %	63	6,3%
Total	**543**	**29,3 %**	**517**	**17,0 %**	**620**	**21,0%**

Table 2 HIV prevalence in CSWs (blood samples, "take all" sampling), Republic of Moldova, 2003 - 2007

Data collection site	2003		2004		2007	
	Sample	HIV prev	Sample	HIV prev	Sample	HIV prev
Chisinau, capital city	**150**	**4,6%**	**151**	**8,5%**	**243**	**2,9%**
Balti	n/a	n/a	n/a	n/a	122	32, 8%
Edinet	n/a	n/a	n/a	n/a	34	2,9 %
Orhei	n/a	n/a	n/a	n/a	69	0 %
Ungheni	n/a	n/a	n/a	n/a	20	25,0 %
Total	**150**	**4,6%**	**151**	**8,5%**	**488**	**10,9 %**

Table 3 HIV prevalence in MSM (blood samples, "take all" sampling), Republic of Moldova, 2003 - 2007

Data collection site	2003		2004		2007	
	Sample	HIV prev	Sample	HIV prev	Sample	HIV prev
Chisinau, capital city	**118**	**1,7 %**	**121**	**2,5%**	**83**	**4,8 %**

In the face of these challenges, estimations of sizes of key populations at risk in Moldova vary according to source and methodology. As most estimates are outdated and of questionable quality, key stakeholders in Moldova have failed to reach a consensus and has waived reporting on sizes of MARPs in the framework of the UA 2008 reporting, pending on more credible estimates.

[3] Used syringes, "take all" sampling
[4] Used syringes, time location sampling
[5] Blood samples, random sampling

Summary of Mission

Meetings

The Consultant met with:

- Ms. Gabriela Ionascu, UNAIDS Country Coordinator for Moldova
- Ms. Alexandrina Iovita, UNAIDS M&E Advisor for Moldova
- Ms Otilia Scutelniciuc, M&E Unit, Moldova National Center for Health Management
- Ms. Svetlana Stefanet and Ms. Larisa Lazarescu, UNICEF Moldova
- Mr. Mihai Oprea, Deputy Director of the National Drug Dispensary
- Ms. Liliana Gherman and Mr Vitalie Slobozian, SOROS Foundation Moldova

The consultant and Ms. Scutelniciuc worked collaboratively to produce size estimations for presentation and discussion at the National Meeting to Plan a New National AIDS Program on February 13, 2010.

This report summarizes the methodologies, initial results, conclusions, and further recommendations coming from these activities.

Size Estimation Methodologies

Capture Recapture

The basic capture recapture methodology begins with a sample from the population of interest. The Lincoln–Petersen method can be used to estimate population size if only two visits are made to the study area, and the assumptions stated above hold (3). Now adopt the following notation:

- *N* = Estimate of total population size
- *M* = Total number of people "captured" and "marked" on the first visit
- *C* = Total number of people "captured" and "marked" on the second visit
- *R* = Number of people captured on the first visit that were then recaptured on the second visit (i.e., included in both samples).

These relationships can be portrayed in a 2x2 table:

		Second Sample		
		Yes	No	
First Sample	Yes	R	b	M = R + b
	No	c	x	
		C = R + c		**N** = R + b + c**+ x**

The relationships from this table can be expressed as:

$$\frac{C}{N} = \frac{R}{M}.$$

Thus, by rearranging, estimated population size is calculated as[6]

$$N = \frac{MC}{R},$$

and this gives an estimate of the total size of the population from which the two sources of data arise.

A modification of the capture-recapture method requiring only one data collection is called the plant-capture method (sometimes called the unique object method). In place of the first sample (capture), study participants whose behavior is indistinguishable from the target population are distributed or planted throughout the population. (Note: this "plant" sample is not randomly selected). Although identifiable after capture, these plants are assumed to be intermixed in the population so that they have a capture probability equal to that of any other population member. As with the traditional capture-recapture model, numbers of the augmented population are captured. The proportion of plants that are caught is an estimate of the probability of capture which, in turn, permits an estimate of population size using a maximum likelihood estimator.

When using the CR method for one sample, several additional assumptions are required. The method assumes that (1) the plants are well mixed into the total population (i.e., the plants are a constant proportion of the target population at each site); (2) the plants accurately identify themselves; and (3) the capture of people from the target population is not affected by the presence of plants (4).This method is analogous or very close operationally to the unique object multiplier method.

The notation used in the plant-capture model is as follows:

---

[6] Confidence interval for the estimate of total population size is given by

$$95\% \, CI = \pm 1.96\sqrt{Var(N)}$$

where Var(N) is calculated as:

$$\frac{MC \bullet (M - R) \bullet (C - R)}{R^3}$$

N the number of individuals in the target population and eligible for enumeration

R the number of planted individuals indistinguishable from target population

P unknown capture probability (for either target population or plant)

n_T the observed number of enumerated members of target population

n_R the observed number of enumerated plants

n $= n_T + n_R =$ total observed or enumerated

Think about the known value R as analogous to the number captured and marked in the first sample of a standard CR model, while n corresponds to the number captured in both samples. The CR formulas above will not apply, since there is a positive probability that n = 0. In this case, we estimate total population size as:

$$N = (R+1)\frac{n_T}{n_R+1} \qquad \text{(A4-3)}$$

And the variance of N is:

$$Var(N) = (R+1)N(1-p+pN)\beta(R+1;p) - N^2(1-(1-p)^{R+1})^2 \qquad \text{(A4-5)}$$

Where

$$\beta(R+1;p) = \sum_{j=1}^{R+1}\frac{1}{j}\binom{R+1}{j}p^j(1-p)^{R+1} \qquad \text{(A-6)}$$

The formula for calculating estimated population size with CR methods is really quite straightforward. However, some important assumptions are implicit in this calculation:

1. The population is closed; that is, the population for the second sampling (capture) exercise includes exactly the same set of individuals as it did for the first (i.e., no in- or out-migration);

2. Individuals captured in both samples can be identified and matched; that is, identifying information is collected in both samples;
3. Capture in the second sample is independent of sample in the first; that is, the chances of someone being included in the first sample are exactly the same as that person's chances of being part of the second sample;
4. Probabilities of capture are homogeneous across all persons and between the two sampling activities; that is, everyone's chance of being included is equal for each sample and between the two samples. This means that both samples must be selected randomly.

Assumption 1 of no movement into or out of the population is easily violated in studies of IDUs or FSWs. If the only change in the population is that some of the members have left, then the method will produce an overestimate of the original larger population (5). The estimate would be biased if, for example, IDUs who are included in the first sample are more likely than others to leave the population by moving away, dying, or ceasing to use. Theoretically, IDUs attending treatment programs may be more likely to reduce their use of drugs for a period, and may be more geographically stable than other users. Movement into the population between the two samples will result in an overestimate of population size. Some biases can be addressed by restricting the study to a relatively brief time period.

Two other related assumptions are that every individual in the population must be equally likely to be included in both samples and that the samples must be independent. If being included in the first sample increases a person's chance of being included in any subsequent samples, the population will be underestimated (5).

<u>Multiplier Methods</u>

The basic principle behind multiplier methods is that the number of people in the population being estimated who appear at a specific institution during a certain time period (e.g., MSM at VCT sites) is equal to the total size of the population of MSM multiplied by the proportion of the population who attended the VCT (6). For example, if the number of MSM who attended VCT in 2007 is known (e.g., from clinic records), and if approximately 10% of IDUs attended treatment in 2007 (from some other source), then the treatment figure can be multiplied by 10 to get an estimate of the size of the IDU population. The method gets its name from the known figure in contact with treatment and the estimated proportion of members of the target population who are in treatment (the treatment multiplier) (7). Then, the estimated population size can be expresses as:

$$S = \frac{P_1}{P_2} * M$$

Network Scale-up Methods

Network scale up method is an indirect technique which helps you to estimate the size of a hard to count population. In this method you ask a random sample of the general population if they know anybody from the population of interest (*E*) such as an IDU. Fundamental to this calculation is the concept of average network size in the general population (called *c*). In implementing this method, it is important to remember:

- c has a distribution, it means that everybody in a population has his/her own *c*, however, we may assume that in average, *c* is constant in each population, and probably, it changes across societies.
- Transmission effects: Everyone knows everything about everyone that they know.
- Barrier effects: Everyone in the general population (*T*) has an equal chance of knowing someone in the sub-population of interest (i.e., in *E*.)
- Inaccurate recall. People don't accurately recall the number of other people they know in the sub-populations that we would ask about.

There are two approaches to estimate *c* in a population; both of these have been validated on populations in the United States; however, little work has been done in international settings.

1. Scaling from known populations: you can ask your samples to mention how many people they know in *E*, as well as in some other sub-populations with known size. For example, you may know the size of a sub-population with aged more than 65, having special names such as "Tatiana", having twins, diabetes and so on. You can estimate *c* based on the responses to the known size sub-populations and use this *c* to estimate the size of unknown population *E* which is illustrated by *e*. However, tutors should help them to understand that the transparency of the sub-populations with known sizes should be comparable with the population of interest, *E*. (This is also called the back-calculation method).

1. Summation method**:** We can directly estimate the network size (*c)* by asking respondents to tell us how many people they know. This is not accurate unless we break it into reasonable sub-groups. We use culturally relevant categories of relation types that are mutually exclusive and comprehensive. These are small enough so that respondents can estimate them reliably.

Consider the following formula:

$$\frac{m}{c} = \frac{e}{t}$$

Where t is the total size of population, e is the size of sub-population E that we would like to estimate its size; on average, each member of our sample knows m people from E, and c is the network size which is the mean number of people known to any member of the general population.
Discuss in your group about the logic behind this very simple formula.

The formula is very simple; e over t shows what proportion of the general population belongs to E; for example you can say that the size of IDUs is 1% of the total population.
m over c shows what proportion of each one's network size belongs to E sub-population. For example, suppose c=300, and in average each sample knows 3 IDUs (m); this ratio is 3/300=1%
Based on the above formula, you can estimate e if you know the other components.[7]

Consider the following formula:

$$(1 - \frac{e}{t})^c$$

This implies the chance of a member of the population who does not know even one member from the sub-population E in case s/he has a network size of c. In other words, it says what the chance is for none of the social network members of a subject belonging to E.

If you understand the logic behind the above formula, you can solve the following equations as well[8]

$$(1 - p_t)^c = (1 - p_r)$$

---

[7] Suppose 0.1 of a population had the history of a car accident. If you select two people randomly, the chance that both of these two people have the history of a car accident is $(0.1)^2=0.01$. On the other hand, you can say that the chance that none of them has the history of car accident is $(1-0.1)^2=0.81$.
Please compute the chance where none of 4 randomly selected people from this population has history of a car accident. This is based on the following relationship:

$$chance\ of\ event\ in\ none\ of\ the\ "n"\ randomly\ selected\ people = (1 - prevalence)^n$$

$where\ p_r$ is the fraction of respondents reporting that they know at least one member of E and p_t is $\frac{e}{t}$.

$$c = \frac{In\,(1-p_r)}{In\,(1-p_t)}$$

$$e = t\left[1-(1-p_r)^{1/c}\right]$$

Moldova Implementation

Moldova Capture-recapture

A key chain was manufactured with a specific logo to make it easily identifiable. The "unique object" has been distributed through outreach workers of harm reduction projects from the data collection sites two weeks prior to the survey. The number of distributed key chains was 20% of the number of IDU-clients of these outreach workers, proportionally to the number of clients that the outreach worker has. A question about whether the respondent has received the unique object in the questionnaire has been used to register the number of respondents with key chains. Since the RDS data collection has not been linked to harm reduction projects, there was a certain degree of independence between distribution of the unique object and participation in the survey.

Moldova Multipler

The study team has also used the multiplier method, in order to attempt estimation of the population size in IDUs in the cities of Chisinau, Balti and Tiraspol. The multiplier co-efficient were determined through specific questions in the questionnaire and matching data with independent data sources benchmarks and by using "unique object" method.

The following existing data sources were used to generate benchmarks:

- HIV registry: number of performed HIV tests attributable to injecting drug use (code XX) in the past year, number of new HIV cases attributable to injecting drug use (code XX)
- Narcology registry: number of registered IDUs, number of IDUs that received detox treatment in year 2008, number of IDUs who have ever received methadone maintenance treatment
- Registries of Harm Reduction projects in Chisinau, Balti and Tiraspol: number of IDUs using syringe exchange services (primary and through

outreach) in the past 12 months, cumulative number of IDUs that have ever used syringe exchange services

The questions in the survey will measure the following multipliers:

- Proportion of IDUs tested for HIV in the past 12 months.
- Proportion of IDUs who received an HIV positive test in the past 5 years
- Proportion of IDUs who are on the registry of the AIDS Center as HIV-infected
- Proportion of IDUs who are on Narcology registry
- Proportion of IDUs who received detox in Narcology Hospitals in the past 5 years
- Proportion of IDUs that used syringe exchange services (NEP and/or outreach) in year 2008
- Proportion of friends of IDUs who have ever used syringe exchange services (NEP and/or outreach)

Muldova Network Scale-up

The Network Scale-Up method has been used in a general population survey performed on a sample of 1969 persons, age group 15 to 64, representative nationally as well as for Chisinau and Balti municipalities, and the North, Central and South region of the country. Moldova used the back estimation method rather than the summation method, as identifying mutually exclusive subcomponents of a personal network may be complicated due to gaps in technical expertise at country level, but also due to different network composition in urban and rural localities of Moldova.

The definition of "known" used derived from the definition used by McCarty et al., with the timeframes limited however to 12 months due to high mobility within the country, as well as high emigration rates. Therefore, egos have been instructed to use the following definition for determining the number of "known" alters:

- people you know and that know you by sight or by name,
- people you could contact
- they have lived in Moldova for the last 12 months
- there has been some contact (either in person, by telephone or mail) in the past 12 months

In order to estimate the number of Injecting Drug Users, Female Commercial Sex Workers and Men having Sex with Men, the questions have been formulated along the following lines:

> “How many X's (separately for IDUs, FCSWs and MSM) do you know and that know you by sight or by name, you could contact them, have lived in Moldova for the last 12 months and there has been some contact (either in person, by telephone or mail) in the past 12 months ?"

The set of known populations have been selected based on a concerted effort to overcome as much as possible the transmission and barrier effects, representing groups of different size and types. Size of the known groups fits within the 0.1 – 4% of the general population that the researchers of the method recommend, and demographic characteristics across groups match as much as possible the demographic characteristics of the general population from which a representative sample has been selected as described above.

The groups of known populations, the size of which is regularly collected by the routine statistics, and the data for which are available for 2008, include:

- women over 70 (4,4 % from the general population)
- men over 70 (2,5%)
- women named Tatiana (3,2%)
- men named Victor (2,2%)
- newborns in 2008 (1,1%)
- men deceased in 2008 (0,6%)
- women deceased in 2008 (0,6%)
- deaths from cancer in 2008 (0,2%)
- deaths due to trauma and accidents in 2008 (0,1%)
- marriages in 2008 (0,7%)
- divorces in 2008 (0,4%)
- kindergarten teachers in 2008 (0,3%)
- graduates of high schools, colleges and gymnasiums in 2008 (1,5%)
- students, higher education in 2008 (3,4%)
- dentists in 2008 (0,04%)
- obstetricians in 2008 (0,02%)
- nurses in 2008 (0,6%)
- policemen in 2008 (0,3%)
- teachers in high schools, colleges and gymnasiums in 2008 (1,1%)
- elderly and handicapped institutionalized in 2008 (0,06%)

To assess the effect of transmission or communication error on estimates of IDU size, Moldova investigators adopted a method of assessing respect first applied in the Ukriane. The NSU survey included an assessment of “respect”; respondents were asked to rate, on a scale of 1 = very low to 5 = very high, various population groups, including the MARPs that were the target of the estimates. Then the data from the respondents were weighted by a factor:

$$W_i = M_i/M_3$$

Where Mi is the average number of IDUs in the network of all people with respect level i for IDUs and M_3 is the average number of IDUs in the network of people with a medium level of respect. Ukraine investigators showed that this addresses bias assuming that people with medium level of respect give most accuarate answers in terms of transmission bias.

Because the Moldovan distribution of respect differed from that in the Ukraine, we made two modifications.

Adjustment 1: If $M_3 = 0$ (i.e., if people with medium level of respect had no IDUs in their network), we added a small amount to the average so the adjustment could be computed. We added 0.5 to M_3; a common correction in statistical analysis.

Adjustment 2: Since the respect questions were on a scale of 1 to 5, we assumed that people with respect of 1 for IDUs report, on average, only 20% of the MARP in their network as people with respect of %. Thus we developed a second weighting:
V_i where:

$$V_1 = 0.6,\ V_2 = 0.8,\ V_3 = 1,\ V_4 = 1.2,\ V_5 = 1.6$$

We hypothesized that network sizes and dynamics might differ between urban and rural areas. Therefore, we calculate estimates separately for the following 8 regions:

- Chisinau
- Balti
- South Rural
- South Urban
- North Rural
- North Urban
- Central Rural
- Central Urban

Results for Moldova IDU

During the consultancy, technical assistance guided final calculations of estimates for the IDU population in selected areas based on the NSU method, and cross-checking those with the multiplier and capture recapture results. The national technical work group shall later on use the same methods to produce estimates of size of SW and MSM, under the overall guidance and with the backstopping of the international consultant.

<u>Capture-recapture</u>

Data for the Capture-Recapture estimate are given in Table 4. Using the formula for the unique object implementation, estimations of population size are calculated.

Table 4. Unique-object Capture-Recapture Survey

Site	**# beneficiaries**	**# keychains distributed (% beneficiaries)**	**Sample Size**	**# respondents with keychain**
Chisinau	1559	234 (15%)	328	3 (0,9%)
Balti	1408	1408 (20%)	370	149 (40,3%)
Tirasplo	270	270 (18,5%)	281	16% (45)

Capture Recapture Estimates are shown as follows

Chisinau

Distribution of key chain	RDS Survey		Total
	Yes	No	
Yes	3	231	234
No	325	x	
Total	328		N

N = 19 094; 95% CI = 2318 – 35 869.

Balti

Distribution of key chain	RDS Survey		Total
	Yes	No	
Yes	149		281
No		x	
Total	370		N

N = 698; 95% CI = 638 – 758.

Tiraspol

Distribution of key chain	RDS Survey		Total
	Yes	No	
Yes	45		50
No		x	
Total	281		N

N= 312; 95% CI = 286 – 338.

The estimate for Chisinau shows poor precision (large confidence interval) due to the small number of key chains returned. These results are influenced by the RDS sampling methodology which may not have produced a sample representative of the population. For these reasons, estimates from these three locations should be interpreted with caution.

<u>Multiplier estimates of IDU Population Size</u>

We used data from two questions from the 2009 BBS for multiplier calculations for Chisinau and Balti:

- Proportion of IDUs that used syringe exchange services (NEP and/or outreach) in year 2008
- Proportion of friends of IDUs who have ever used syringe exchange services (NEP and/or outreach)

We used benchmark data from harm reduction programs and the Mondova Narcology registry for 2009.

Data for 2008	City		
	Chisinau	Balti	Tiraspol
Harm Reduction	1654	1656	270
Syringe exchange multiplier		34,6%	19,1%
Estimate		4731	1421
Friends multiplier	4,7	2,5	4,2
Estimate	7774	4140	1134
Narcology data	2229	469	
multipler	64%	46,5%	36,7%
Estimate	3483	998	

Network Scale-up Estimates of Population Size

<table>
<tr><th>Area</th><th>Average Personal Network Size</th><th>Crude NSU (no adjustment)</th><th>NSU Adjustment 1</th><th>NSU Adjustment 2</th><th>Narcology</th></tr>
<tr><td>Chisinau</td><td>249</td><td>1116</td><td>7896</td><td></td><td>2229</td></tr>
<tr><td>Balti</td><td>207</td><td>466</td><td>9320</td><td></td><td></td></tr>
<tr><td>South Rural</td><td>368</td><td>84</td><td>110</td><td></td><td rowspan="2">155</td></tr>
<tr><td>South Urban</td><td>383</td><td>57</td><td>137</td><td></td></tr>
<tr><td>North Rural</td><td>374</td><td>552</td><td>142</td><td></td><td rowspan="2">172</td></tr>
<tr><td>North Urban</td><td>378</td><td>240</td><td>550</td><td></td></tr>
<tr><td>Central Rural</td><td>455</td><td>71</td><td>71</td><td>101</td><td>212</td></tr>
<tr><td>Central Urban</td><td>189</td><td>1721 (737)*</td><td>62,000 (19,000)*</td><td></td><td></td></tr>
</table>

*One outlier reporting more than 1500 people in network was excluded.

Conclusions

The most valid size estimates come from a comparison of estimates from several methods. We illustrate this comparison with estimates for the IDU population in Chisinau.

Estimator	Data Source	Estimate
Official Statistics	Narcology	2220
Capture-Recapture	Unique object, RDS	8548 – 42 620
Multiplier	Narcology	3483
Multiplier	Harm reduction/friends	7774
Multiplier	Harm reduction/needle exchange	12 626
NSU unadjusted	NSU survey	1116
NSU Adjustment1	NSU survey	7896

For Chisinau, we see consistent results between the multiplier estimate and the scale-up estimate adjusted for transmission error: 7774 to 7896. While the Capture-recapture estimate was larger, the lower confidence limit is near the range of other estimates, and the large variability is due to the small number of key chains returned in the RDS sample. The narcology data, even when combined with a multiplier, gives a low number and does not seem to be useful. The multiplier estimates are subject to limitations as well. For example, registration may be based on official place of residence rather than place of enrollment in the survey. This might cause the multiplier estimates to be underestimates of the true population size.

We can conclude that a reasonable estimate for the size of the IDU population in Chisinau (2008) could be taken as 8000.

Limitations

First, not all risk groups were considered. We did not analyze data for for prisoners (although official statistics are available for this population), street children/most-at-risk adolescents, and migrants.

Second, Narcology data available for this consultancy were not useful for estimation of MARA due to concerns that inconsistencies in data reported for the same time periods may affect data quality.

Third, the Ukranian adjustment for transmission error may not be appropriate for Moldova. Experience with the alternative adjust described earlier should prove useful.

Considerable effort was expended by staff in providing harm reduction data in a useful format. Although not a limitation, these data are very valuable and should be collected and reported in a more consequential manner

Finally, we did not incorporate age of respondent as a variable in our analysis. We will consider such an extension in further work to improve the validity of size estimation.

Next Steps

1. Complete analysis of NSU estimates with Adjustment 2.
2. Complete analyses for additional multipliers from BSS 2008.
3. Incorporate Age of respondents in NSU analysis (See appendix A).
4. Investigate appropriateness of known populationsion (See Appendix B).

5. Rerun all NSU analyses on the basis of adjustment and known population decision.

6. Review estimates of other MARPS (e.g., partners of IDUs, commercial sex workers, MSM), and compare estimates from capture-recapture, multiplier methods, and survey results.

The Consultant will work remotely with M&E Staff (via e-mail and Skype) to accomplish these analyses between 4 March and 10 March 2010. The outcome of the Consultancy will be an estimated national range for sizes of MARPs, as specified by MOH.

Further Recommendations

The following recommendations follow from activities undertaken in this mission. These are beyond the scope of the current contract but should be considered by Moldova M&E staff.

1. Evaluate the feasibility of having harm reduction data available in a more usable, accessible format. For example, cumulative numbers should be converted to unique clients for each year

2. Narcology data from the Drug Dispensary were not useful for this size estimation project. Evaluate the quality of these data and consider feasibility of increasing quality and timeliness for size estimation purposes.

3. Any size estimation results from Moldova should be published so that other countries can benefit from the experience and so that Moldova can be seen as a leader in the region in this type of public health analysis. Publication could be in regional newsletters or in peer-reviewed journals (See www.journals.lss.com/jaids/pages/default.aspx. or www.oxfordjournals.org/oup_journals/ije/about.html.) Such methodological publication or dissemination should ensure no adverse reaction against beneficiaries.

4. Consider separate data collection for MARA. Questions regarding personal network size and known populations will differ for

adolescents. In addition, ethical considerations give special protections to adolescents in data collection.

5. Evaluate the feasibility of collection of data for prisoners and migrant populations.

6. Routinely evaluate data collection activities to ensure their utility for program delivery and intervention (8).

The presentation based on this mission is attached.

Appendix A: Adjusting NSU for Age of Respondent

It is possible that age of the respondent may influence the number of IDUs in a personal network, a form of barrier effect. Thus, we propose to analyze the data in age strata within the urban/rural areas presented. If there are J age strata, then let E_i = NSU adjusted estimate for stratum i and N_i = number of respondents in stratum i. Then the overall age-adjusted NSU estimate will be:

$$E = \frac{\sum_{i=1}^{J} N_i x E_i}{\sum_{i=1}^{J} N_i}.$$

The standard deviation of E is given by:

$$StDev(E) = \sqrt{\sum_{i=1}^{J} N_i^2 s_i^2}$$

Where s_i^2 is the variance of E_i. A 95% confidence interval can then be constructed as E $\pm$ 2*StDev (E).

Appendix B: Selecting Known Populations

Known populations should first be selected based on availability of official statistics. Once that is satisfied, known populations should be represented in the data in approximately the same proportion as they exist in the population. Thus, to select appropriate known populations, we can compare the distribution of known population members reported by respondents to that distribution for the official statistics.

This can be accomplished through calculating frequencies for both the sample and official statistics and making a comparison. Suppose that SAMPLE contains the number of network members from a particular known population group, and OFFICIAL represents the official statistics for that group. Following are several approaches using SPSS.

Frequency Counts Using Dot Plots

Dot Plots can be used to provide a simple frequency chart for SAMPLE and OFFICIAL thus enabling us to compare the resulting distributions to each other. Creating these simple frequency charts is best done using the **Interactive Graphs** menu in SPSS.

To begin, simply click on the **Graphs** menu and choose **Interactive**. In that sub menu, choose **Dot**. This kind of chart is called a "dotplot", in which each dot represents a number (or several numbers, if they happen to be the same) in the distribution. When the **Create Dots** panel appears, you will see a list of all the variables (the classes in our case) in the left window and spaces representing the vertical and horizontal axes in the right window.

Simply highlight the data that you want to examine (SAMPLE, in this case), and drag it into the box for the horizontal axis. When you have done that, click **OK**, and the chart will pop up in the **Output** window. The window also provides a number of options for adding titles, lines, and various other accessories to your chart.

Now do the same thing for OFFICIAL. (Notice that when you choose another variable to plot, it automatically replaces the one in the horizontal axis box.) Fortunately, SPSS makes it extremely easy for you to examine these charts and compare them. To do so, simply locate the small icon in the left panel that represents each of the charts. When you single-click on one, you will see it in the **Output** window.

Numerical Summaries of Distributions

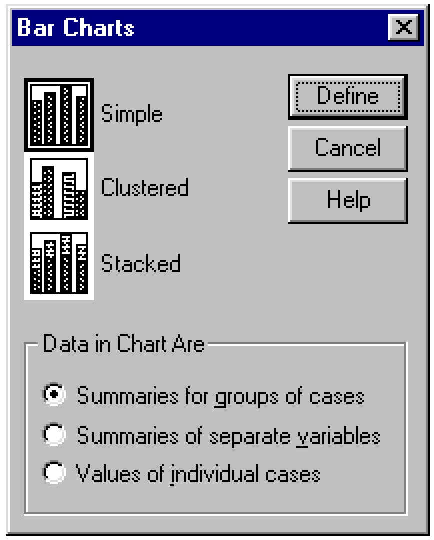
To produce numerical summaries of any distribution just go to the **Analyze** menu, choose **Descriptive Statistics**, and then choose **Frequencies**. When you get the **Frequencies** dialogue box, simply highlight each variable name, and then click on the arrow to put it in the right side of the box. You can do the variables one at a time or all at once. Then click on the box labeled **Statistics**. This will take you to a dialogue box labeled **Frequencies: Statistics**, where you can decide which descriptive statistics you would like to have calculated.. When you selected the statistics you want, click **Continue**. Then click **Ok**, and SPSS will calculate the descriptive statistics.

Bar Charts

One of the best ways to present data is by using a bar chart or a histogram. Bar charts are appropriate when you have discrete sets of data (males, females) while histograms should be used when you have a continuous variable (numbers of people known). Below are instructions on how to produce each type using *SPSS*.

In our situation, we have data on numbers of population members in both our sample and the official statistics. The easiest way to compare graphically is to create a bar chart. To do that, follow the following steps.

First, select **Bar…** under the **Graphs** menu. When the **Bar Charts** panel comes up, you will see that you have several selections of the kind of bar chart you want. Start with **Simple**.



There are a few things to point out about the SPSS output file that you have created. The information comes in an **Output** window, and it has two parts. On the right is the output from whatever procedures you have just completed. On the left is the **Outline** – an outline representation of all the recent SPSS procedures that you have run. You can explore this panel as you want, but the main thing to remember is that it can be a good way of helping to keep track of the procedures that you have run. You can also edit, rearrange, and delete items in your output file using this feature.

REFERENCES

1 Info Bulletin on the Epidemiologic situation in HIV/AIDS, Implementation of the National HIV Program, and Targets for 2009, NCPM, 05.03.09.

2 Scutelniciuc & Bivol, 2008

3 Millar T, Comingo-Salvany A, Eastwood C, Hay G. Glossary of terms relating to capture-recapture methods. J Epidemiol and Comm Health 2008;62(8):677-81.

4 Goudie IBJ, Jupp PE, Ashbridge J. Plant-capture estimation of the size of a homogeneous population. Biometrika 2007;94(1):243-8.

5 Sanderland RL. Estimation of the number of heroin users in New South Wales using police arrest data: development of a statistical model. Research Grant Report Series B 86/1. New South Wales Drug and Alcohol authority, United Kingdom.

6 Walker N, Stover J, Stanecki K, et al. The WORKBOOK approach to making estimates and projecting future scenarios of HIV/AIDS in countries with low level and concentrated epidemics. Sex Transm Infect 2004;80(Suppl 1):i10–3.

7 Zhang D, Wang L, Lv F, et al. Advantages and challenges of using census and multiplier methods to estimate the number of female sex workers in a Chinese city. AIDS Care 2007;19:17-9.

8 See www.cdc.gov/surveillanceevaluation. Accessed 16 February 2010.