



**Economic and Social
Council**

Distr.
GENERAL

ECE/CES/GE.30/2006/7
20 June 2006

Original: ENGLISH
ENGLISH AND RUSSIAN

ECONOMIC COMMISSION FOR EUROPE

STATISTICAL COMMISSION

CONFERENCE OF EUROPEAN STATISTICIANS

Group of Experts on Gender Statistics

Fourth session

Geneva, 11-13 September 2006

Item 3 of the provisional agenda

**TOWARDS INTERNATIONAL STANDARDS FOR DATA COLLECTION AND
STATISTICS ON VIOLENCE AGAINST WOMEN**

Invited paper submitted by Lancaster University, United Kingdom*

I. ABSTRACT

1. In the context of the tremendous and rapid developments at national and international level in the development of survey methodology for the collection of data on violence against women, it is timely to consider the development of international standards. There are two key elements: first, the indicators for which the data is being collected; secondly, the survey methodology necessary to produce such statistical data in an internationally comparative form in the context of a mainstreaming strategy. The indicators need to be simple, meaningful, robust and relevant to a range of international settings. They should measure both extent and severity, including prevalence, number of incidents per population unit, and level of injury. The

* This paper has been prepared by Sylvia Walby, Lancaster University, UK at the invitation of the secretariat.

methodology needs to include a comprehensive sampling frame that does not exclude marginal and disadvantaged groups; the conditions for disclosure of sensitive events, such as self-completion; and an adequate sample size.

II. INTRODUCTION

2. There is very considerable experience in the development of survey methodology for collecting data on violence against women. National surveys have now been carried out in many countries in both the North and South (Alm  ras et al 2004; Johnson 1996; Tjaden and Thoennes 2000; Walby and Allen 2004), as well as developments in comparative work using the same instrument in groups of countries (Garcia-Moreno et al 2005; Kishor et al 2004). There has been time to reflect on these developments, to consider the implications of different ways of asking questions, so as to build on and move beyond early formulations.

3. In several areas alternative strategies have emerged. This paper will review these alternatives and make recommendations as to how to proceed. They include: the determination of summary indicators that focus on the specificity of the violence or the use of mainstream categories, such as the range of actions specified and the range of perpetrators included; establishing a sampling frame and response rate so as to ensure the most comprehensive coverage including of disadvantaged and marginalised groups in the context of limited resources; the location of the questionnaire as either a stand alone survey or a module of another survey; the method of delivery of the questionnaire so as to maximise response and disclosure.

III. INDICATORS AND DEFINITIONS

4. The selection of the indicators for which supporting information is to be collected is the first step in devising a survey. It is not an after-thought or a side issue, since this will shape the range of questions to be asked as well as affecting decisions on the comprehensiveness of the sampling frame, as well as the importance of response rate and disclosure of particularly sensitive matters.

Why indicators?

5. Indicators summarise complex data into a form that is meaningful for policy makers. They constitute a key link between an evidence base and policy making. There have been many policy innovations to reduce and eliminate violence against women; much political good-will; and much rhetoric. In order to decide whether initiatives are having a positive impact it is necessary to know whether the situation is deteriorating or improving. There are many forms and types of knowledge about the nature of violence against women and the policies to stop this. Often these data are too complicated to support the decision-making of policy makers without the input of considerable time and expertise. The purpose of indicators is that they provide a simple summary of a complex picture, abstracting and presenting in a clear manner the most important features needed to support decision-making.

Criteria for selecting indicators

6. Several criteria for the selection of the indicators have been developed (Berger-Schmitt and Jankowitsch 1999; Alméras et al 2004; Statistics Canada 2002). In general, indicators should:

- (a) Summarise complex data;
- (b) Be unambiguous and easy to interpret;
- (c) Enable an assessment as to whether an improvement or deterioration has occurred including the establishment of whether changes over time have occurred;
- (d) Be meaningful and relevant to policy makers, service providers and the wider informed public;
- (e) Be capable of being supported by reliable and robust quantitative data;
- (f) Be available at regular intervals and be comparable between countries and population groups;
- (g) Be neither so many as to confuse, nor so few as to mislead.

7. There are several proposals for indicators, including the European Union (2004); European Women's Lobby (2003); Statistics Canada (2002); US Center for Disease Control, National Center for Injury Prevention and Control (Saltzman et al 2002); UK government (Home Office 2005); UN Economic Commission for Latin America and the Caribbean (Alméras et al 2004); United Nations High Commissioner for Refugees (2003); summarised in Walby (2006).

8. There are two major kinds of indicators: outcomes; and policy actions. The focus here is only on outcomes, since the measurement of the extent of the development and implementation of policy requires a different methodology from population surveys.

IV. DEFINITIONS

9. The development of indicators requires decisions on how to define the range of actions in a manner suitable for a large scale survey in a wide range of countries and the range of perpetrators.

10. The United Nations (1993) definition of gender-based violence in the 1993 Declaration on the Elimination of Violence Against Women is: 'Any act of gender-based violence that results in, or is likely to result in, physical, sexual or mental harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or private life'.

11. This UN definition is the starting point, but it needs more precise operationalisation before it is suitable for inclusion as a set of questions in a survey. It is possible to separately name forms of violence against women, including domestic violence, sexual assault, stalking, sexual harassment in the workplace, female genital mutilation, forced marriage, dowry deaths and so-called 'honour' crimes. However, in some countries some of these forms of violence would be present in very few cases in the survey, possibly making it preferable to group them together into more general categories, even if the questionnaire asks separate questions about different forms.

12. There is a productive tension in the development of indicators between on the one hand the detailed appreciation of distinctive nuances within a framework specific to the expert VAW field, and on the other hand, the use of more general categories that facilitate addressing priorities within mainstream policy arenas. Early attempts at specifying indicators tended to the former, while later developments have tended to move towards the latter.

V. INDICATORS OF EXTENT AND SEVERITY OF GENDER-BASED VIOLENCE

13. There are two dimensions to outcome measures:

- (h) Extent: measured by either or both of prevalence and the number of incidents;
- (i) Severity: measured by one or more of frequency, the nature of the action and injury.

Extent

14. There are two main approaches to the operationalisation of the extent of gender-based violence: prevalence and incidents.

15. Prevalence refers to the proportion of the population that has experienced violence in a given period, usually either (adult) life-time or the previous year. The notion of prevalence captures the particular and specialised nature of domestic violence as a coercive 'course of conduct', a series of related occurrences, rather than a one-off event. This figure, which ranges from around one quarter to one half of women in their life-time, has been important in the establishment of the scale of the problem (Krug et al 2002).

16. A disadvantage of the use of prevalence as the sole indicator of the extent of VAW is that it is a specific measure developed in the specific field of VAW, which can be hard to mainstream into some other policy domains, especially that of crime, where the focus is instead on the number of incidents.

17. A further disadvantage of prevalence as the sole indicator is that it does not contain a measure of severity. In those countries where surveys of inter-personal violence are addressed to men as well as to women, the focus on prevalence can sometimes obscure the extent of gender inequality in the use of violence. If domestic violence enters crime statistics as a 'course of conduct', then it counts as just one crime incident, even though there are usually several events within this 'course of conduct'. In this way, the repetition and frequency of the attacks disappears from view thereby leading to underestimates of the extent of violent crime and domestic violent crime in particular.

18. A count of the number of incidents of VAW in a population unit is an alternative or additional indicator of the extent of gender based violence. The use of this indicator facilitates the mainstreaming of VAW into other policy domains, such as crime. In operationalising this, care needs to be taken to distinguish between the number of actions and the number of events.

19. This requires rethinking the use of the Conflict Tactics Scale is being used, since it is hard to clearly separate the number of actions from the number of events in this framework (Fals-Stewart, Birchler and Kelly 2003).

20. It is recommended that the indicators for the extent of gender-based violence should include both prevalence and the number of incidents, not prevalence alone.

Severity

21. It is important to have an indicator that represents the severity of the violence, as well as its extent. There are three main approaches to the measurement of severity: frequency; the nature of the action; whether there is an injury and if so its seriousness.

22. Frequency. The number of times that the same person is subject to violence is a measure of severity. This is different from the number of incidents per population unit, being rather the average number of incidents per person subject to the violence.

23. Action. The nature of the action has very frequently been used as a measure of its seriousness, especially through the use of the Conflict Tactics Scale (CTS), developed by Straus and Gelles (1990). The CTS consists of a list of items, of increasing levels of severity, such as slap, kick, choke, use a weapon. There are three problems with this scale. First it is confined not only to domestic violence, but also excludes sexual violence in a domestic context, so it is far from comprehensive. Additional scales would need to be used to capture sexual assault and non-intimate partner violence. Second, the use of the actions of the perpetrator as the measure of severity may not be congruent with its effects on the victim. In particular, the injurious effect of the same action is greater when the perpetrator is a man and the victim a woman than vice versa (Walby and Allen 2004). Third, the scale is unique to the field of domestic violence, and it is hard to use it to mainstream VAW into other policy fields which use different forms of scaling of severity, such as injury.

24. Injury. The existence and severity of injury is a further way of operationalising the concept of severity in gender-based violence. This is a victim-focused measure of the severity of the impact of the violence. The existence or not of any physical injury is the easier part of such a scale. A further distinction may be made as to the levels of violence, using a commonly agreed scale. Such an approach has the advantage of being easy to mainstream into adjacent policy fields such as the criminal justice system and health. To be comprehensive there would need to be a concept of injury which goes beyond that of simple physical injury. For example rape is a very serious injury in its own right, not needing to be further translated. A more difficult, though important issue, is that of the mental health injuries that are a common consequence of VAW, but which are simple to operationalise for a population survey. More complex surveys might include measures developed in the mental health field, but this probably goes beyond reasonable expectations for international standards for a VAW population survey.

25. Recommendations: Severity should be indicated in two ways. First, the frequency of the incidents. Second, by injury, including no injury, minor injury, severe injury, rape.

VI. CONSISTENCY

26. There are a number of matters concerning time periods and the specific population studies that are important if there is to be comparability between survey outcomes in different surveys. These include: time-period during which the violence occurred; age of population sample; marital and cohabitation status of population sample.

Time period

27. The time period used in most VAW surveys has been that of either or both of the last year, life-time (or adult life time), though there have been some exceptions with the use of periods of five years, and of six months.

28. The use of life time was useful in the early stages of development of this field, especially for consciousness raising. However, this time period is not of much relevance if the focus is on the evaluation of the impact of recent policy changes.

29. The use of quite short periods for recall, such as that of six months, may aid the accuracy of recall. However, the shorter the period the larger needs to be the sample size, so this may be not practical in the context of budget restraints.

30. The time period of 'last year' is the most commonly adopted approach. This is the time period recommended here.

Age

31. Surveys are quite varied in the use of age restrictions on the population sample. Most have a lower cut off that approaches adulthood, though this varies, predominantly between 15 and 18. Many but not all have an upper cut off. The criteria range from reproductive age of 49, to ability to use a computer unaided by interviewer of 59, with further ages representing various expectations of competence.

32. There needs to be agreement on a key age range for which data is collected and reported, even if individual countries have their own reasons for including younger or older groups.

33. The recommendation here is 16-59 as the core age group.

Marital and Cohabiting Status

34. There is restriction in some surveys, especially where the focus is domestic violence rather than the full range of gender-based violence, to women who are currently or ever have been married or cohabiting. This is unduly restrictive, since violence against women can take place outside of marriage and cohabitation.

35. The recommendation is that no restrictions of current or previous marriage or cohabitation are placed on the population sample.

VII. SURVEY METHODOLOGY

36. During the development of national surveys of gender-based violence, there has been much reflection on and refinement of methodological dilemmas and challenges. These include: free-standing survey or add-on module; the achievement of a comprehensive sampling frame; the best method of delivery of the questionnaire; and whether confidentiality or rapport with the interviewer is more likely to facilitate disclosure of sensitive events.

Sample Size

37. In order to be able to measure different forms of violence, not only physical domestic violence, but others, such as rape, it is necessary to have a sample size large as possible, ideally not less than 10,000 persons.

Free-standing or add-on module?

38. Data on gender-based violence is positioned in surveys in three ways. One practice is to include a few questions within a mainstream survey. A second is to conduct a free-standing surveys in its own right. A third is to add a relatively self-contained module onto a more mainstream survey. There is a tension here between creating the best possible survey environment and the need to keep costs down.

39. The inclusion of a few questions on VAW within an already existing survey has the advantage of least cost, but the disadvantage of collecting a restricted amount of information in an interview context that has priorities other than VAW.

40. The advantage of a dedicated survey is that it can be tailor made for the methodological needs of investigating violence against women; the disadvantage is the practical one that the resource base for an annual dedicated survey is less likely to be available than for a survey that also gathers information on mainstream matters.

41. There are a number of issues about interviewing where the priorities of a generic and a dedicated survey may diverge. For instance, an interviewer who is both female and specially trained has an advantage in eliciting responses in such sensitive areas, however, the prioritisation of such practices in interviewing is less likely to be achieved when there is a generic rather than dedicated survey. Further, a private context for the interview results in higher rates of disclosure than when there is someone else present in the room (Walby and Allen 2004). The achievement of this private context for the interview may be more likely to be prioritised in the context of a dedicated than a generic survey, but there is no intrinsic reason why a generic survey context could not deliver this.

42. A compromise solution is the use of a relatively self-contained module on gender-based violence, which is attached to a mainline survey on a related topic. This uses an existing well-resourced annual national survey that has an established base of expertise and resources and attaches to this a self-contained, specially introduced, specialised module of relevant questions. The cost of the additional questions is modest as compared with the establishment of a dedicated annual survey. The UK BCS has experimented with this form (Walby and Allen 2004) and is now committed to regularly attaching special modules on domestic violence, sexual assault and stalking (Home Office 2005).

43. While the ideal is the dedicated survey, it is likely that cost restrictions would prevent regular repetition. Thus, it is probable that a self-contained module on gender based violence, attached to a related survey may be the most pragmatic solution.

Sampling Frame and Response Rate

44. The ideal sampling frame is one that includes all members of the population. This is most closely approached in the census, but many other national surveys may fall short of this ideal to varying degrees, often for reasons of time and resources. The ideal response rate is very high. A comprehensive sampling frame and a high response rate are probably of greater importance in surveys of violence against women than in many other surveys, since it is likely that those who fall outside the sampling frame or are not reached or do not respond are more likely to have been subject to violence than those who have not. It is the more 'marginal', excluded and disadvantaged groups of women who are most likely to have been subject to violence, especially in the near past, and these are precisely the groups that are most likely to be omitted if short cuts or economies are taken with the development of the sampling frame and survey instrument. While for many other types of surveys the omission of this section of the population from the sampling frame may not be considered sufficiently important to be worth the expense and effort to include them, for surveys on violence against women this is a potentially significant omission.

45. For example, women who have fled to refuges, to temporary residence with friends and kin, to emergency bed and breakfast or hostel accommodation, or who are homeless in the immediate aftermath of a domestic assault are most likely to be omitted from sampling frames and to have low levels of response to the survey. Samples based on women who have gone to refuges and shelters have consistently shown much higher rates of frequency of abuse than those from national surveys (Dobash and Dobash, 1979; Straus and Gelles 1990). The omission of the most heavily abused section of the population is a problem for a survey attempting comprehensive coverage and accurate estimates. This is a significant omission for the measurement of domestic violence in the last 12 months, although it may have less impact on the life-time rate of domestic violence since some women may now be living in settled violence free homes.

46. There are ways of supplementing the sampling frame to include these populations, which could enhance future surveys. These include drawing up additional sampling frames based on lists of hostels, refuges, and other temporary accommodation that could be provided by those who fund and run such accommodation. In addition, the procedure for sampling the person in residential households could include all who are actually staying there, not merely those who are permanently resident. However, this is hard to achieve and no VAW survey has yet managed this.

47. The recommendation is that considerable priority is given to obtaining a comprehensive sampling frame and a very high response rate, while recognising that this can be expensive.

Mode of delivery of questionnaire

48. Surveys have been carried out using: postal questionnaires, telephone, face-to-face interviewing, and by self-completion on a computer. While some suggest that there is little evidence that it makes much difference, others have argued for particular methods, especially either telephones, or telephone or face-to-face, or for self-completion by computer (Walby and Myhill 2001).

49. Postal questionnaires usually have the lowest response rate of all methods, so are usually considered inappropriate for those surveys where this is important, as is the case in of surveys of violence against women. However, Statistics Finland used a postal questionnaire and obtained a surprisingly high response rate of 70% (Heiskanen and Piipisa, 1998). This might be explained in terms of the unique features of Nordic society.

50. In some countries landline telephones have been used to make contact with respondents. Whether this provides a comprehensive sampling frame depends on two major factors that vary between countries. In some countries, such as Canada, coverage is nearly comprehensive, while in others such as the UK, telephone ownership rates in private households in Britain are particularly low among the poor heads of lone adult households (Beerton and Martin, 1999), and these are likely to include disproportionate numbers of women who have fled a violent home.

51. A further problem with telephone surveys today is the use of mobile phones for which random dialling techniques are not available, thus excluding those who have only mobile phones from the sampling frame. These are likely to be disproportionately young, probably leading to a skewing of the results, since younger people tend to have higher rates of violence.

52. The mode of enquiry is relevant because of its potential to not only reduce the response rate but to omit key sections of the population that have been found to be most at risk. The best approach will vary between countries at different levels of development of communication infrastructure. It may that sampling house is the best way to obtain a comprehensive sampling frame, although this has implications for cost.

Self-completion: Rapport or confidentiality?

53. Is rapport or confidentiality more conducive to disclosure of events that may be sensitive? On the one hand there is the possibility that face-to-face interviewing can build up more rapport and support disclosure of sensitive events, while on the other hand, confidentiality engendered by strategies such as self-completion by computer or by postal questionnaire may increase the likelihood of respondents divulging sensitive information.

54. The British Crime Survey provided an opportunity to compare the outcome from face-to-face interviewing with that of the confidentiality of self-completion. In the first part of the interview, face-to-face interviewing is used; in the second, there is self-completion of specialised modules, with the computer handed over to the respondent. The respondent reads the questions from the computer screen and enters their responses into the computer. Only the respondent can see the questions and the answers they have given. The prevalence of domestic violence is five times higher when the more confidential self-completion methodology is used, as compared with the more traditional face-to-face interviewing. While there are other differences between the two questionnaires that may contribute to the differences, nonetheless, the scale of the difference suggests that confidentiality is more important than rapport in facilitating the disclosure of domestic violence (Walby and Allen 2004).

55. One disadvantage of this method is that it depends upon literacy, which may be unevenly distributed within some populations. The model described here requires laptop computer- based

interviewing, which, while increasingly common in developed countries is expensive, though it may be that other forms of self-completion may also have similar effects.

56. The recommendation is to deliver the questionnaire in a manner that prioritises confidentiality in order to facilitate disclosure.

VIII. CONCLUSIONS

57. The following are recommended as guidelines for the mainstreaming of data collection efforts on violence against women.

58. Indicators: Indicators of gender based violence should cover two dimensions: extent and severity. Extent is best captured by both the prevalence of gender-based violence in the population and by the number of incidents per population unit. The number of incidents is needed in order to facilitate the mainstreaming of gender-based violence into adjacent policy domains, such as the criminal justice system. The severity of the violence is best captured by both the frequency of incidents per abused person and by the level of injury. An indicator for severity is particularly important in those countries which use the same instrument to address violence against men as well as violence as well women in order to reveal the gender asymmetry of the violence.

59. Consistency: There needs to be consistency on time-periods, age cut-offs, and other statuses in order to enable comparability. It is recommended that the best period for recall is the last year, that 16-59 be the core age group, and that no restrictions on marital or cohabiting status be imposed.

60. Survey methodology: A large sample size is best, especially to allow for the separate identification of less common forms of violence such as rape. If there is sufficient finance, then a free-standing survey is better than one that is integrated into another survey. But a relatively autonomous module added onto a related survey may offer the best balance between appropriate framing and cost effectiveness. It is very important that the sampling frame is comprehensive and does not omit marginal or disadvantaged groups. Although there is increasing use of telephone surveys in developed countries, it may be that the best way to achieve this in most countries is through sampling people in houses. The provision of confidentiality to the respondent is very important in facilitating disclosure.

BIBLIOGRAPHY

Alméras, D., R.Bravo, V. Milosavljevic, S. Montaña and M. Nieves Rico (2004) *Violence Against Women in Couples: Latin America and the Caribbean*. (Santiago, Chile: UN Economic Commission for Latin America and the Caribbean).

Berger-Schmitt, R. and Jankowitsch, B. (1999) *Systems of Social Indicators and Social Reporting*: EUReporting WP No.1. Mannheim: Centre for Survey Research.

Dobash, R.P., Dobash, R.E., Wilson, M. and Daly, M. (1992) 'The myth of sexual symmetry in marital violence', *Social Problems*, 39, 1:71-91.

European Union (2004) *Council Resolution on the review of the implementation of the Beijing Platform for Action*. 1502/04. SOC 558. JAI 497. Brussels.

European Women's Lobby (EWL) (2003) *Proposal for a Policy Framework and Indicators*. (Brussels: European Women's Lobby).

Fals-Stewart, William, Gary R. Birchler and Michelle L. Kelly (2003) 'The timeline followback spousal violence interview to assess physical aggression between intimate partners: Reliability and validity', *Journal of Family Violence*, 18, 3, 131-142.

Garcia-Moreno, C., C.Watts, H.Jansen, M. Ellsberg and L.Heise (2005) *WHO Multicountry Study on Women's Health and Domestic Violence against Women*.

Heiskanen, M. and Piispa, M. (1998) *Faith, Hope, Battering - A Survey of Men's Violence against Women in Finland*. (Yliopistopaino, Helsinki: Statistics Finland).

Home Office (2005) *Domestic Violence: A National Report* (Home Office)
<http://www.crimereduction.gov.uk/domesticviolence51.pdf>

Johnson, H. (1996) *Dangerous Domains - Violence Against Women in Canada*. (Canada: Nelson Canada).

Kishor, Sunita and Kiersten Johnson (2004) *Profiling Domestic Violence: A Multi-Country Study*. (Calverton MD: Measure DHS+ORC Macro).

Krug, Etienne G., Linda L. Dahlberg, James A. Mercy, Anthony B. Zwi and Rafael Lozano (eds) (2002) *World Report on Violence and Health*. (WHO).

Saltzman, Linda E., Fanslow, Janet L., McMahon, Pamela M., and Shelley, Gene A. (1999) *Intimate Partner Violence Surveillance Uniform Definitions*. (National Center for Injury Prevention and Control, Centers for Disease Control: Atlanta, USA).

Statistics Canada (2002) *Assessing Violence Against Women: A Statistical Profile*. (Ottawa: Statistics Canada).

Straus, M.A. and Gelles, R.J. (eds) (1990) *Physical Violence In American Families*. (New Brunswick, New Jersey: Transaction Publishers).

Tjaden, Patricia and Nancy Thoennes (2000) *Full Report of the Prevalence, Incidence and Consequences of Violence Against Women*. NCJ 193781. (US Department of Justice, National Institute of Justice: Washington DC, USA).

United Nations (1993) *Declaration on the Elimination of Violence Against Women*. UN General Assembly 20 December 1993. UN Doc. A/RES/48/104.

UNHCR (2003) *Sexual and Gender-Based Violence against Refugees, Returnees and Internally Displaced Persons*. (Geneva: UNHCR).

Walby, Sylvia (2005) 'Improving the statistics on violence against women', *Statistical Journal of the United Nations Economic Commission for Europe*, 23, 3-4.

Walby, Sylvia (2006) *Developing Indicators of Gender-Based Violence*. Mimeo.

Walby, Sylvia and Jonathan Allen (2004) *Domestic Violence, Sexual Assault and Stalking: Findings from the British Crime Survey*. Home Office Research Study 276. (London: Home Office). <http://www.homeoffice.gov.uk/rds/pdfs04/hors276.pdf>

Walby, Sylvia and Andrew Myhill (2001) 'Comparing the methodology of the new national surveys of violence against women', *British Journal of Criminology*, 41,3, 502-22.
