Annex 4

Technical considerations in the use of spirometry

This annex describes recommended procedures in the recording, performance and interpretation of spirometry in programmes of screening and surveillance of workers exposed to mineral dust.

**Recommended procedures and quality control**

Spirometry tests should be conducted in accordance with the American Thoracic Society’s (ATS) recommended spirometry standards, as well as the recommendations of the European Respiratory Society (ERS) (Quanjer et al., 1993; ATS, 1995). These standards not only establish minimum requirements for the equipment used, but also the procedures to follow in administering tests. A quality control programme is a critical component of spirometry screening. The quality control programme should include the adoption of a procedure manual describing the proper calibration, use and maintenance of all equipment, requirements for record maintenance and the procedures of technician training and monitoring. When screening information is collected from multiple sites, centralized review of test quality is needed. If spirometry results are to be interpreted longitudinally (i.e. if several tests performed at various times on the same individuals are interpreted), the quality control centre should attempt to identify survey biases. A survey bias is an unexplained change in a group’s mean FEV₁ between surveys. A record of calibration tests is particularly useful when a survey bias is suspected, so that instrumentation errors can be evaluated as a source of bias.

**Interpretation**

**Test reproducibility**

The first step in interpreting spirometry is to assess the quality of the test. The lack of a sufficient number of acceptable test trials or inconsis-
tency in results should be carefully considered during interpretation. The presence of excessive flow oscillations in the spirogram, resulting in the identification of the tracing as potentially unsatisfactory because of presumed cough, may instead indicate a functional or structural disorder. The lack of a reproducible test result may be caused by disease and has been shown to be associated with an increased risk of mortality in cohorts that are occupationally exposed to pathogens (Eisen et al., 1985; Kellie et al., 1987). In addition, shorter individuals may have more difficulty in meeting reproducibility criteria than taller individuals. Therefore, an individual’s results may sometimes be safely interpreted even though the test is not considered reproducible by ATS standards (ATS, 1995).

Comparison with reference values

The recommended method of interpreting a single spirometry observation involves the comparison of an individual’s observed FEV₁ values with a reference value derived from cross-sectional data that take into account subjects’ height, sex and age. Both ATS and ERS have recently published statements on the interpretation of spirometry results (ATS, 1991; Quanjer et al., 1993). The cut-off values selected to separate individuals for whom no intervention is warranted (presumed “normal”) from those for whom a preventive intervention is recommended or required (presumed “abnormal”) should be chosen to reflect the goals of the screening programme. For purposes of screening, where the early identification of abnormality is the goal, these cut-off values may be different from those generally used in clinical practice, where the focus is on disease diagnosis and confirmation. Test sensitivity (the ability of a test to identify accurately the presence of disease), specificity (the ability of a test to identify accurately the absence of disease) and the predictive value of both positive and negative results vary depending on the cut-off values adopted and on the prevalence of disease in the screened population.

Selection of reference values

Reference values for spirometry testing should be selected on the basis of methodological, epidemiological and statistical criteria. Reference values are generally derived from regression equations generated from data on lung function gathered in healthy (often non-smoking) populations. Published reference values vary, not only for technical reasons, but also because of differences in the population mean of the groups
studied. These differences may relate to socioeconomic, psychosocial and other factors. The ATS, for example, does not recommend a universal reference value for all populations, but instead recommends that, to the extent possible, reference values be based on values obtained using comparable equipment in a population with comparable age, physical characteristics, socioeconomic background and ethnic origin (ATS, 1991). By contrast, the ERS guidelines recommend the use of one equation for males and one for females, based on pooled data collected from several countries. Almost all reference values are based on, at a minimum, an individual’s age, sex and height.

For ethnic groups where specific reference values may not be available, some adjustment of the values obtained in Caucasian populations may be possible. For example, the ERS recommends that, for subjects of African descent, predicted values be multiplied by 0.87. This procedure is not recommended by the ATS, nor does it appear to be justified on the basis of a recent analysis of published data from over 30000 men and women of sub-Saharan African descent (White et al., 1994). Some variability between ethnic groups may be due to differences in the average trunk length relative to average standing height (Quanjer et al., 1993).

**Criteria for abnormal FEV₁, based on comparison of reference values**

Although the 95th percentile (the value above which 95% of the population scores) is often used as the lower limit of normal (LLN) for purposes of clinical interpretation, this may not be appropriate for screening and surveillance. In some circumstances, where the purposes of cross-sectional screening would be better met by a more sensitive indication of abnormality, the 85th or 90th percentile might be selected as the level that triggers further monitoring, investigation or other action. The LLN is available, or can be calculated, from the data published for most reference values. For example, the ERS notes that a LLN approximating the 95th percentile can be estimated by subtracting 0.84 litres from the predicted FEV₁ value in men and 0.62 litres from the predicted FEV₁ value in women.

**Criteria for abnormal FEV₁, based on changes over time**

A comparison of an individual’s current FEV₁ value with his or her own FEV₁ value from the past may be useful, particularly for workers whose FEV₁ is above the predicted value derived from a reference population. A quality control programme is, it should be reiterated, especially im-
important if longitudinal changes are to be assessed. Because of considerable variability in FEV₁ values over the short term, a year-to-year change of less than 15% should not be considered significant. For periods of observation longer than a year, adjustment for the expected annual decline in FEV₁ is appropriate. Therefore, the LLN for a follow-up FEV₁ can be computed by taking 85% of the baseline value minus the expected decline over the period. An individual’s expected decline depends on his or her age, but for practical purposes a value of 25 ml per year is often recommended. For example, an individual whose initial FEV₁ is 4.00 litres would be considered to have an accelerated decline in FEV₁ if his or her FEV₁ fell below 3.15 litres in 10 years ((0.85 × 4.00) − (10 × 0.025)). This approach has been presented in more detail recently (Hankinson & Wagner, 1993).

To increase the sensitivity of spirometry for screening purposes, comparisons of an individual’s FVC and FEV₁/FVC with the appropriate reference values can also be done. However, because the FVC is usually a more difficult parameter to determine accurately (being more dependent on effort than the FEV₁), the FVC and FEV₁/FVC comparisons should be used only when there is reasonable assurance of their reliability.

References


Annex 5
Questionnaire development and use

Questionnaires are often employed in evaluations of the effects of worker exposure to mineral dusts. An increased level of reported symptoms in the workplace should trigger and focus an environmental investigation. The Medical Research Council in the United Kingdom (MRC), American Thoracic Society, European Coal and Steel Community and International Union Against Tuberculosis and Lung Diseases have published respiratory symptom questionnaires for use in epidemiological studies (Ferris, 1978; MRC, 1986; Burney & Chinn, 1987). Programmes for the screening of workers exposed to mineral dust often draw their questions from these more comprehensive research tools. Critical issues in the design and use of respiratory symptom questionnaires have been reviewed in detail (Samet, 1978; Attfield, 1986). This annex summarizes the relevant guidance for questionnaire use in the screening and surveillance of workers exposed to mineral dust.

Questionnaire screening in isolation from other screening methods lacks utility. Questionnaires should also be a part of a wider prevention programme. They aid in the identification of respiratory symptoms, trigger further medical investigation and treatment or support workplace interventions to decrease exposure for workers. Conditions caused by mineral dust exposure are reviewed briefly in Chapter 5 of the text. Of the conditions listed, chronic bronchitis alone is defined by responses to questions. Other conditions commonly have symptoms that can be detected by questionnaires, but for most, more definitive diagnostic methods are available. At the other extreme, there is currently no direct medical benefit from using questionnaires to screen for work-related malignancies, because of the short period between the onset of symptoms and the time that medical care is sought, as well as the lack of effective interventions. Nevertheless, the documentation of malignancies can be important in surveillance programmes.

Unlike other medical tests, questionnaires gather information that the worker already knows. The information derived from question-
naires can be very useful in permitting observation of patterns of symptoms in the workplace, but workers will generally require an explanation of the importance of both individual and group results. To ensure comprehensive and successful screening, it is essential to include workers and their representatives in all phases of the development and administration of questionnaires and in the interpretation and reporting of their results, as well as in the design of plans based on these results.

The questionnaire may be self-administered or administered by a trained interviewer, depending on local conditions. The use of interviewers provides the benefit of consistent questionnaire administration and may be necessary in dealing with workforces comprising several language groups or including illiterate workers. The possible disadvantages of using interviewers include the time needed to train the interviewers and potential interviewer bias. A self-administered questionnaire has the benefit of reducing training requirements and personnel time but has the potential drawback of introducing a non-response bias if workers ignore some questions. To avoid a non-response bias, self-administered questionnaires may be checked for accuracy and completeness while the respondent is waiting. The advantages and

<table>
<thead>
<tr>
<th>Table 1. Comparison of the methods of questionnaire administration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interviewer-administered questionnaires</strong></td>
</tr>
<tr>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td>• Structured administration</td>
</tr>
<tr>
<td>• Reliability</td>
</tr>
<tr>
<td>• Adaptable to several language groups or to illiterate workers</td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
</tr>
<tr>
<td>• Time and cost of training interviewers</td>
</tr>
<tr>
<td>• Possible observer bias</td>
</tr>
<tr>
<td>• Inconvenience (i.e. need to schedule appointments, etc.)</td>
</tr>
<tr>
<td><strong>Self-administered questionnaires</strong></td>
</tr>
<tr>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td>• Minimal cost and personnel requirements</td>
</tr>
<tr>
<td>• Convenience (i.e. can be mailed, etc.)</td>
</tr>
<tr>
<td>• Less stressful for workers (i.e. can be completed when desired and without real or implied pressure)</td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
</tr>
<tr>
<td>• Possible non-response bias</td>
</tr>
<tr>
<td>• Person other than the designated respondent may complete the questionnaire</td>
</tr>
<tr>
<td>• Must be simpler than an interviewer-administered questionnaire</td>
</tr>
<tr>
<td>• Requires literate respondent or surrogate</td>
</tr>
<tr>
<td>• Multiple versions necessary for workforces comprising several language groups</td>
</tr>
</tbody>
</table>
disadvantages of interviewer-administered questionnaires and self-administered questionnaires are summarized in Table 1.

Steps in questionnaire design and development

*Clarify the goals of the programme*

- Will the questionnaire be used for screening, surveillance or both?
- What are the diseases or conditions of concern?
- What kind of information would trigger a response or intervention?

*Identify existing resources*

- Is there already a questionnaire or section of a questionnaire reported to be useful in screening for the conditions of interest?
- Can existing questions be translated into the relevant languages and format?

*Determine notification and reporting requirements*

- What information will be needed for efficient notification of workers?
- What, if any, information will be aggregated, analysed and reported outside the worksite (either voluntarily or compulsorily)?

*Plan for data analysis*

- Have the personnel who will be analysing data been involved in the process of questionnaire design?
- Will a particular format facilitate data analysis?
- Will data entry or analysis be technologically aided (e.g. by use of optical scanning technology)?

*Select question form and wording*

- Clear, specific, short questions often stimulate the most reproducible results and are easiest to analyse and track over time.
- “Skip patterns” are used in most standard questionnaires to elicit more information about a positive answer. If the answer is negative, the respondent “skips” to the next topic.
- Will the questions be understood by the respondents? Is the
language used appropriate for their educational and cultural background?

- The word-for-word translation of questions from an existing questionnaire may lead to misunderstanding, and it is usual to translate and reverse-translate questionnaires several times before they are considered ready for use in the target language. If comparability of results with other workforces elsewhere is an objective, questions should, if possible, be identical to those used in the other workplaces.

- Worker representatives and others familiar with the local language should assist in the translation and design of questionnaires. The ATS questionnaire has produced useful, valid, and comparable results in English- and French-speaking workforces in Quebec (Osterman et al., 1991), as have the IUATLD questionnaires in English-, French- and German-speaking populations in Europe (Burney et al., 1989).

- In many developing countries, which typically have experienced the pressures of rapid urbanization, workforces are made up of several different language groups, and the practice of translation and back-translation in three or more languages is not practical. Even though the reproducibility of responses may be adversely affected if workers are not interviewed in their mother tongue, free translation by multilingual interviewers from a standard questionnaire (e.g. the ATS questionnaire) can give answers to certain questions that are comparable to those obtained in more homogeneous workforces (Becklake et al., 1987).

**Determine frequency of questionnaire administration**

Depending on the characteristics of the target disease and the conditions of exposure, parts of questionnaires may be usefully administered with different frequencies. In many instances, a comprehensive baseline questionnaire is useful on entry into the workforce. Subsequently, only periodic information on work exposure, medical history and symptoms need be recorded. Since questionnaires are not recommended as an exclusive means of screening workers exposed to mineral dust, the frequency of questionnaire administration is generally determined by the periodicity of other programme components.

**Organize the questionnaire logically**

Group together questions that elicit demographic information, symptoms of concern, relevant medical history, the history of past work
exposure, current work and exposures, use of protective equipment and other exposure history (e.g. at home or during leisure time).

**Edit the questionnaire**

- Assess the relevance and importance of each question.
- Ensure that the minimum data needed to fulfil the goals of the programme will be collected.

**Write instructions for questionnaire administration and coding**

Develop a coding system that is consistent with data analysis. If a computer will be used for data entry, allot areas on the questionnaire for a coding system that meets the requirements of the software.

**Pre-test the questionnaire**

- Administer the questionnaire to a limited number of workers and involve them in determining whether the questionnaire functions as intended. Assess the clarity, validity and reproducibility of the questionnaire and identify any sources of bias.
- Revise the questionnaire as necessary.

**Questionnaire analysis**

The use and analysis of the questionnaire results should be planned when questions are selected and intervention strategies are developed. Actions to be recommended on the basis of questionnaire results should also be defined. When data will be grouped and reported, or analysed over time for an individual, the following steps should be included in data analysis:

- Checking questionnaires for completeness.
- Maintaining a logging or tracking system for each questionnaire.
- Coding questionnaires for data entry.
- Entering data on paper forms or into a computer.
- Verifying the accuracy of data entry.
- “Cleaning” data by checking for outliers and inconsistencies. (A computer can print out the distribution of every variable. Values that fall outside an acceptable range should be checked.)
- Creating computer files for statistical analysis (when appropriate).
• Reporting results to participants.
• Conducting and reporting analysis of aggregate data.

Intervention

Interventions resulting from questionnaires may include referral for testing (e.g. spirometry or chest X-ray), referral for medical examination or more comprehensive environmental monitoring and exposure control. Upon completion of the aggregate analysis, exposure and risk groups can be identified. When risk groups are recognized, controls should be implemented to limit further exposure and unfavourable health effects.

Assessment, review and revision

The contribution of the questionnaire to achieving the goals of the screening programme should be periodically assessed. The questionnaire should be revised when needed, with the understanding that changes in the wording, ordering and length of the questionnaire may produce results that are not comparable over time. Nevertheless, the non-comparability of results is of less concern in screening than in the conduct of surveillance programmes or longitudinal epidemiological research. However, unclear questions and questions producing no useful information should always be revised or eliminated. If possible, new questions should first be used along with the original questions in order to check their comparability before the original questions are dropped. If new information about health risks becomes available, more questions should be added.

References


SCREENING AND SURVEILLANCE OF WORKERS EXPOSED TO MINERAL DUST


