



HAL
open science

French permanent survey on indoor air quality - Part 2.: Questionnaires and validation procedure of collected data

Mickael Derbez, Anthony Gregoire, Olivier Ramalho, Julien Garrigue,
Séverine Kirchner

► To cite this version:

Mickael Derbez, Anthony Gregoire, Olivier Ramalho, Julien Garrigue, Séverine Kirchner. French permanent survey on indoor air quality - Part 2.: Questionnaires and validation procedure of collected data. *Healthy Building'2006*, Jun 2006, Lisboa, Portugal. pp.327-331. hal-00688541

HAL Id: hal-00688541

<https://hal.science/hal-00688541>

Submitted on 17 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

French Permanent Survey on Indoor Air Quality – Part 2 : Questionnaires and Validation Procedure of Collected Data

M. Derbez¹, A. Gregoire¹, O Ramalho¹, J. Garrigue¹ and S. Kirchner¹

¹Centre Scientifique et Technique du Bâtiment, 84, avenue Jean Jaurés – Champs-sur-Marne
F-77447 Marne-la-Vallée Cedex 2, France
email: mickael.derbez@cstb.fr http://www.air-interieur.org

Summary: *The questionnaire data collected during the first national campaign on indoor air exposure inside dwellings must be validated previous to data exploitation. This article describes the three successive stages of the validation procedure currently applied in the database, i.e. control of database homogeneity, management of missing data and coherence check. Preliminary results are presented and show the efficiency and simplicity of the method. At the issue of the validation process, the study of the representativeness of the investigated population is realized and exploitation of validated measurement data can begin.*

Keywords: questionnaires, field study, validation procedure

Category: Indoor air exposure

1 Introduction

The French permanent survey on indoor air quality conducts field measurement and questionnaire campaigns in different indoor environments, e.g. dwellings, offices, schools... Each campaign has a specific objective and as such needs specific development in survey strategy, questionnaires and measurement protocols.

The first national campaign focused on dwellings was conducted from October 2003 to December 2005. A representative sample of 567 dwellings from the French national housing stock –approximately 24 million main homes– was randomly selected [1] and investigated. The objectives of this campaign are first to provide with target pollutant levels inside dwellings, second to identify sources and determinants of indoor air pollution and finally to assess occupants' exposure.

The data collected by questionnaires will be used to study the representativeness of the investigated population with the reference population, to identify the environmental and behavioral determinants of indoor air pollutants and to reveal groups of differently exposed inhabitants according to their behavior and socio-demographic characteristics.

However, the data collected by questionnaires must be validated before conducting data analysis.

After a presentation of questionnaires used in the survey, the paper describes the control and validation procedure of questionnaire data and details preliminary validation results.

2 Methodology

2.1 Presentation of questionnaires

Standardized questionnaires elaborated and tested during a pilot survey conducted on 90 dwellings from March to July 2001 [2] have been optimized for the national campaign focused on dwellings.

Four categories of questionnaires are used throughout the campaign in addition to the measurement of defined key pollutants. Measurement and analytical protocols are presented elsewhere [3]. Most of questionnaires present closed questions with single or multiple answers. Quality code can be used in the following cases: no answer, not asked question, refusal to answer.

The **carbon monoxide questionnaire** is aimed at the detection of poisoning risk from carbon monoxide exposure in dwellings. A risk exists in the following cases:

- technicians identify defects in the internal combustion installation of the dwelling during the survey,
- indoor CO level is higher than 60 ppm during the survey week,
- exhaled CO level is higher than 15 ppm for non-smoker or 35 ppm for smoker,

The **description questionnaire** is used to detail the characteristics of both dwellings and household inhabitants. This questionnaire consisted of nine parts:

- a first contact questionnaire in order to collect practical information for the survey organization.
- a sensory questionnaire to assess initial perceived air quality (completed by the technician).
- a household questionnaire that describes outdoor environment, building construction characteristics,

standard equipment inside dwelling, living and cleaning habits of occupants,...

- a room description questionnaire that lists for each room, e.g. nature of heating energy, ventilation.

system, wall, floor and ceiling coverings, furniture,...

- a bedding questionnaire providing information on age and size of bed and mattress where dust mite allergens are measured.

- an individual questionnaire describing all individuals in the household (child caring mode for children less than 10 years, profession and nature of professional exposure, domestic activities, use of cosmetics, tobacco exposure for inhabitants over 10) and gathering results from exhaled CO measurements on voluntary individuals over 6.

- a retrospective questionnaire to list all particular events that occurred during the survey week (use of heating system, cleaning products, air fresheners, cleaning and cooking activities,...)

- a survey quality questionnaire to know the householder opinion on the quality of investigation the duration of survey,...

- a household survey form that includes dates of initial and final visit, number of individual questionnaires distributed and the drawn day of daily diary completion.

The third category of documents, **time-activity diaries**, is left to the inhabitants during the week of investigation. Each children and adult must complete

two time-activity diaries, one for the survey week (weekly diary) and one for a randomly chosen day of the week (daily diary). Both paper questionnaires are completed to inquire spending time and room location of the inhabitants throughout the dwelling every 10 min. The daily diary details the used room, the activities, the number of occupants at a given time (smoker or not) and the products used.

The fourth category of questionnaire focuses on **allergy and respiratory symptoms** for each inhabitant at least 15 years old.

2.2 Collect of information by questionnaires

Collect of information by questionnaires is made by technicians who were formed in the same way by CSTB. Six training courses of 4 or 5 days were organized and 50 technicians were formed on the use of survey tools (both questionnaires and measurement protocols of key indoor pollutants). All questionnaires are detailed and technical terms of building or dwelling facilities are described and illustrated.

Three groups of questionnaires are used in the survey (table 1):

- **face to face interview** (FTFI) to collect detailed data when the technician directly communicates with the reference person or each inhabitant,

- **self-administered questionnaire** (SAQ) filled either by the reference occupant or by other individuals,

Table 1. Detailed characteristics of questionnaires

Questionnaire name	mode of administration	moment of completion	nb of items / questionnaire	nb of items / survey	approximate duration of completion (mm:ss)
Carbon monoxide	TFQ	after survey	6	6	00:30
Description					
first contact	FTFI	before survey	24	24	06:00
sensory	TFQ	first visit	5	10	01:00
household	FTFI	first visit	267+10	267 + 10 x i	50:00+06:00 x i
room description	TFQ	first visit	87	87 x j	05:00 x j
bedding	TFQ	first visit	12	12	01:00
individual					
<10 years	FTFI	first visit	23	23 x k	08:00 x k
10 years or over	FTFI	first visit	51	51 x (i-k)	16:00 x (i-k)
retrospective	FTFI	second visit	56	56	12:00
survey quality	SAQ	after survey	14	14	03:00
household survey	TFQ	first visit	8	8	01:00
Time-activity diary	SAQ	during survey	-	-	-
Health	SAQ	after survey	50	50 x m	04:00 x m

i : household inhabitant number

j : dwelling room number

k : household less than 10 years inhabitant number

i-k : household aged 10 or over inhabitant number

m : household aged 15 or over inhabitant number

nb: number

- **technician filled questionnaire (TFQ)** to acquire objective data of the building or dwelling characteristics by visual inspections.

Each technician used a mobile application, running on personal digital assistant (PDA). This application allows the technicians to collect data through an electronic questionnaire. At the end of each survey, technicians can dump their data into the main database by connecting to the extranet website [4]. The technician can check that information has been correctly imported into the main database. If electronic questionnaires are modified, surveyors can update the application via the extranet website. Most of questionnaires are electronic and only four stay in paper format (first contact questionnaire, survey quality questionnaire, time-activity diary, health questionnaire). At the moment, only first contact questionnaires have been computerized and sent to database.

Collect of information is mainly made during the survey. At the first visit which lasts approximately 3h30, one of both technicians (tech. 1) completes the survey questionnaires in presence of the reference occupant. The second technician (tech. 2) carries out all room inspection, performs immediate sampling and measurements and sets up continuous measuring instrument.

During 7 days, inhabitants must fill time-activity diaries. After the survey and at the second visit (1h30), retrospective questionnaire is administered and measuring instrument are removed. Time activity diaries are checked, completed and given back to the technicians.

The number of items per survey and the approximate length of time required to complete each questionnaire depend on the number and age of the household occupants and on the number of rooms in the dwelling. For the FTFI questionnaire, the completion duration takes into account the time needed to clarify technical terms, to answer the questions of the inhabitants, etc... Number of items and completion duration for time activity diaries are not indicated since too few items are inquired.

Let us take the example of a household with 2 adults and 2 children of less than 10 years (i=4, k=2, i-k=2, m=2) living in a dwelling with a living room, a kitchen, two bedrooms, a bathroom, a independent water closet (j=6) and without taking into account time activity diaries.

The total number of items approximately reaches 1200 and is fairly divided between questions asked to occupants (633) and those asked to technicians (568). The majority of questions addressed to the occupants (82 %) is administered by face to face interview.

The time required to complete all questionnaires is estimated to last approximately 3h30. The completion of FTFI questionnaires during the first visit is the longest (approximately 2h20) while 35 minutes are enough for questionnaires filled by technicians.

3 Presentation of the validation procedure of questionnaires data and preliminary results

This control and validation procedure of questionnaire data aims to ensure the quality of collected data and the coherence of the database. The results presented here are temporary because control and validation procedure are still in process. At the moment, all electronic questionnaires and only one paper questionnaire (first contact questionnaire) have been checked. Other paper questionnaires will soon be integrated into database.

The control procedure allows to check the number of questionnaires according to the number of household occupants and dwelling rooms. This control is undertaken immediately after the uploading of data by the technician towards the database (via extranet website) and ensures the recovering of missing questionnaires from technicians.

The validation procedure is performed in **three points**.

First of all, it is necessary to **control the database homogeneity by** tracking all empty, double and missing questionnaires. For 16428 questionnaires integrated in database, practically all questionnaires (99.9%) are complete and only 14 questionnaires remain empty or double.

From the 16414 complete questionnaires, the average response rate is high (97.8%) but overestimates the final response rate as it does not take into account SAQ (survey quality questionnaire, time-activity diary, health questionnaire). As assumed, this rate differs according to the questionnaire type. It is higher for TFQ (99.6%) than FTFI (96.2%).

Table 2 presents the response rate of each questionnaire and shows that the response rate is lower than 98% for three questionnaires.

Table 2. Number and response rate of complete questionnaires

	questionnaires nb	% response rate
Carbon monoxide Description	567	99.5
first contact	567	85.1
household survey	567	99.3
sensory	1130	99.9
household	6148	98.9
room description	4690	99.7
bedding	567	98.2
individual		
<10 years	251	77.5
10 years or over	1361	91.1
retrospective	566	99.0

For the first contact questionnaire (85.1% response rate), the technicians judged the questionnaire not useful enough as it deals with survey organization (number of building floors, dwelling floor, presence

of an elevator, balcony ...). For individual questionnaire (77.5% for < 10 years and 91.1% for 10 years or over), several questions relative to the exhaled CO are filled by quality code instead of a measured value. Measurement of the exhaled CO must be realised in both visits but it is difficult to gather all household occupants twice. Furthermore, the large difference (13.6%) between the response rates of individual questionnaires is due to a lower number of items in the questionnaires intended to children up to 10 years old.

The second validation point deals with the **management of data missing**. Data missing could be recovered from other data and logical rules. More rarely and only for important questions, it is possible to ask once more the technician or the investigated occupant. Statistical methods can also be used in order to estimate missing data.

This point is illustrated in the case of a question relative to smoker/non-smoker status that has not been readily asked to every occupant aged 10 or over. This question is important in order to identify the determinants of indoor air key pollutants and, unfortunately, the associated response rate was unsatisfactory.

Children of less than 10 years were assumed not to smoke. For the other individuals, the smoker/non-smoker status was not known for 362 (27%) out of 1361 individuals (Fig. 1). A total of 267 answers (20%) was recovered with the use of questionnaires and logical rules. About 95 answers (7%) are still missing but may be recovered with collected data from the health questionnaire and time activity diaries.

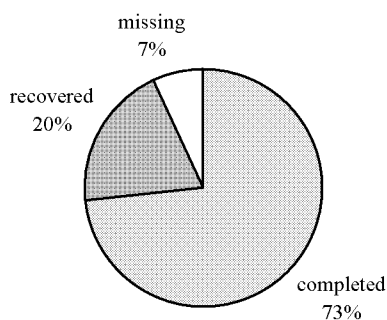


Fig. 1. Completion rate for the smoker/non-smoker status item.

The information inquiry began by gathering all the answers to questions relative to tobacco consumption from every individual questionnaire. These questions describe for each occupant, the type of smoker (regular, occasional smoker, ex or non-smoker), the number of cigarettes smoked in the 24 last hours and the hour of the last smoked cigarette.

In the case it is insufficient, the information can be found out from the logical association of answers to various questions (first contact, household and room

questionnaire). For example, the answer of question "household smoker number" allows to conclude that nobody smokes when the answer is zero. In the opposite case, we have to take into account the other validated variables like number of household occupants, number of non-smokers, number of regular smokers, etc.

The third validation point permits the **check of data coherence**. Constraints exist for some answered questions, e.g. unit of measurement, dwelling area/room number ratio. By means of redundant information search throughout the database and elaboration of automated logical rules, incorrect answers (differing measurement unit or different answers to a same question) can be replaced. To date, almost 200 logical rules were elaborated. Decision rules must be established (identification of the most reliable variable, concordant answers among inquired occupants, answer of household reference occupant more reliable as regards another inhabitant, technician answer represents the major weigh). These decision rules depend on the item and are specific to each case.

Ventilation system is an interesting example since, on the one hand, it can explain indoor air key pollutants levels and, on the other hand, few households really know how to describe and correctly identify their own ventilation system. This information is confirmed by looking more closely at the presence of continuous mechanical extract ventilation (CMEV) system in dwelling. The answer given by the reference person in household questionnaire (what is the nature of your ventilation system ?) is compared to the information taken by the second technician (tech. 2) in room description questionnaire. This system is characterized by the presence of air inlet in the main rooms (bedroom, living), air outlets in service rooms (kitchen, bathroom), and exhaust airflow by ducts.

According to the database, 247 dwellings seemed to possess a CMEV but the analysis of coherence (Fig. 2) shows that this information is only proven for 194 dwellings (79 %). In 38 dwellings (15%), technicians do not agree with the reference person (technician recognized the CMEV) and in 15 dwellings (6%), the opposite is observed (technician didn't recognise the CMEV).

In this example, we decided to privilege the answer of the technician rather than the household answer because the technician is formed and thus has more knowledge and experiment. Furthermore we checked the information given by the technician in room description questionnaire by other means like technician remarks after survey, combination of dwelling rooms information, or information collected during measurement of the exhaust flowrate in service rooms. As a result, 232 dwellings (94%) are equipped with a CMEV and 13 dwellings (5%) with another system. The system of ventilation could not be correctly recognized in two dwellings (1%).

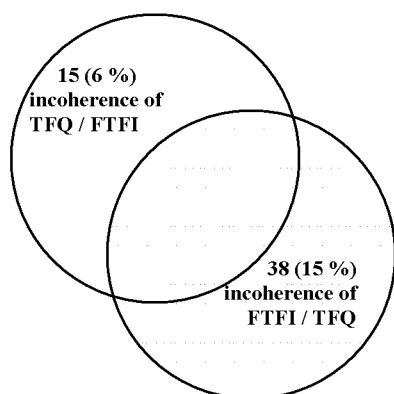


Fig. 2. Validation of the presence of a continuous mechanical extract ventilation (CMEV) system in dwelling by coherence between FTFI questionnaire data and TFQ data.

The check of data coherence can be made on a combination of variables like the dwelling area/number of occupants ratio, the dwelling area/number of rooms ratio, etc.

4 Conclusion and discussion

The data collected by questionnaire within the first national campaign focused on dwellings are in the course of validation. The preliminary results of validation of the data show that practically all the questionnaires are complete, that the response rate is high and that it is possible with the aid of logical rules, to recover the missing data and to test the coherence of the database.

These results does not take into account the analysis of self-administered questionnaires who will soon be integrated into the database. For time activity diaries, the control of missing data is only possible during the collection of diaries by technician after the survey week. Technician can recover missing data by interviewing occupants during the second visit. For each individual, the concordance of declared presence in a room will be checked between daily and weekly diaries. For survey quality questionnaires and health questionnaires, the control of missing data and the coherence check are more difficult because the questionnaires are sent after the second visit. Moreover, the answers to the questions remain strictly personal and confidential for each household occupant (perceived survey satisfaction, health).

After the validation of questionnaires data, the first task is to study the sample representativeness of the 567 investigated dwellings as regards the national reference (general census of population of 1999 conducted by INSEE). A sample weighting procedure will be used to adjust the investigated sample.

Data analysis will follow sample adjustment and will need the creation of aggregated variable from existing variables.

Acknowledgements

This French permanent survey on indoor air quality is currently receiving financial support from the French Ministry in charge of Construction and Housing, the French Ministry in charge of Health, the French Ministry in charge of Environment, the Agence de l'Environnement et de la Maitrise de l'Energie (ADEME) and the Centre Scientifique et Technique du Bâtiment (CSTB).

References

- [1] F Golliot *et al.* The French national survey on indoor air quality: Sample survey design, *Proc. Healthy Buildings 2003*, December 7-11th, 2003, Singapore, pp 712-717.
- [2] S Kirchner *et al.* The French permanent survey on indoor air quality – Survey design in dwellings and schools, *Proc. Indoor Air 2002*, June 30 – July 5, 2002, Monterey, California, pp 349-354.
- [3] O Ramalho *et al.* French permanent survey on indoor air quality – Part. 1: Measurement protocols and quality control. *Healthy Buildings 2006*, submitted.
- [4] N Bus *et al.* Information system for indoor air quality studies. *Proc. Indoor Air'2005*, September 4-9th, 2005, Beijing, China, pp 138-143.

