

Study population, questionnaire, data management and sample description

Chiara de Waure, Andrea Poscia, Andrea Viridis, Maria Luisa Di Pietro and Walter Ricciardi

Istituto di Sanità Pubblica, Sezione di Igiene, Università Cattolica del Sacro Cuore, Rome, Italy

Abstract

Introduction. This article describes methodological issues of the “Sportello Salute Giovani” project (“Youth Health Information Desk”), a multicenter study aimed at assessing the health status and attitudes and behaviours of university students in Italy.

Materials and methods. The questionnaire used to carry out the study was adapted from the Italian health behaviours in school-aged children (HBSC) project and consisted of 93 items addressing: demographics; nutritional habits and status; physical activity; lifestyles; reproductive and preconception health; health and satisfaction of life; attitudes and behaviours toward academic study and new technologies. The questionnaire was administered to a pool of 12 000 students from 18 to 30 years of age who voluntarily decided to participate during classes held at different Italian faculties or at the three “Sportello Salute Giovani” centers which were established in the three sites of the Università Cattolica del Sacro Cuore (Catholic University of the Sacred Heart of Rome).

Results. The final study sample was composed by 8516 university students. The mean age of responders was 22.2 (Standard Deviation 2.0) and 5702 (67.0%) were females. According to the distribution in age classes, 3601 (43.3%) belonged to the 18-21 one, 3796 (44.5%) to the 22-24 class and 1019 (12.2%) to the 25-30 class. With respect to socio-economic status, data were available for 8410 responders and showed that 50.3% of students belonged to the middle class.

Discussion. The project took into consideration a large number of individuals from different regions of the country and therefore may be considered representative of the general population of Italian university students. Furthermore, it is the first to address, at the same time, several issues, in particular attitudes and behaviours toward health, in Italian university students.

Conclusion. The analysis of data from such a large sample of university students sets the basis for identifying the most appropriate interventions in order to address the specific needs of this population.

Key words

- university students
- Italy
- demographics

INTRODUCTION

In order to be able to interpret results from every kind of studies and assess their generalizability, it is necessary to provide readers with information on study population, tools used and methods applied to manage data.

This article describes methodological issues of the “Sportello Salute Giovani” project, a multicenter study aimed at assessing the health status and attitudes and behaviours of university students [1].

The article describes the study population, the development process of the questionnaire used to carry out the study, the questionnaire contents, the data entry and cleaning, the codification of variables and the sample characteristics.

STUDY POPULATION

The study population was represented by university students aged from 18 to 30 years attending the faculties listed in *Table 1*.

Questionnaires were given to a pool of 12 000 students who voluntarily decided to participate. Students were made aware of the possibility to enroll in the survey during classes and at the three “Sportello Salute Giovani” which were established in three sites (Rome, Milan, Brescia) of the Università Cattolica del Sacro Cuore (Catholic University of the Sacred Heart).

QUESTIONNAIRE

The questionnaire was adapted for our study population from the Italian health behaviours in school-aged

Table 1
Universities involved in the project

| University | Region |
|---------------------------------------|--------------------------|
| Università Cattolica del Sacro Cuore | Lazio, Lombardia, Molise |
| Università degli Studi di Milano | Lombardia |
| Sapienza Università di Roma | Lazio |
| Università degli Studi di Tor Vergata | Lazio |
| Università degli Studi di Roma Tre | Lazio |
| LUMSA | Lazio |
| Università degli Studi di Padova | Veneto |
| Università degli Studi di Cagliari | Sardegna |
| Università degli Studi di Messina | Sicilia |
| Università Politecnica delle Marche | Marche |

children (HBSC) questionnaire [2]. In particular, the HBSC was adapted by a group of 22 experts in the field of epidemiology, fertility and reproductive health, nutrition and eating disorders, sports medicine, bioethics, physical and mental health, forensic and legal medicine. Furthermore, the adapted version was validated with experts of the Istituto Superiore di Sanità (Italian National Institute of Health) which funded the study.

The final questionnaire was composed of 93 items and addressed the following dimensions:

- demographics;
- nutritional habits and status;
- physical activity;
- lifestyles;
- reproductive and preconception health;
- health and satisfaction of life;
- attitudes and behaviors toward academic study and new technologies.

Closed ended and multiple choices questions were used to address the abovementioned dimensions.

The questionnaire was anonymous and self-administered to students who gave back it in special boxes.

The Italian version of the questionnaire is available online as Supplementary Material at www.iss.it/anna.

DATA MANAGEMENT AND SAMPLE DESCRIPTION

Data collected through questionnaires were entered into databases through a dedicated mask, built on Epi-Info, by 11 collaborators. The 11 collaborators were trained in a 6-day course given by the experts involved in the project. The course addressed both university students' health issues, counselling and data management.

The 11 databases were merged together by means of SPSS version 22, but records missing at least one whole dimension of the questionnaire were deleted.

Final database consisted of 8734 observations. Some new variables were also created. The student's age was calculated as the difference between the year of the end of the study (2013) and the birth year. The Body Mass Index (BMI) was obtained dividing the weight (kg) by

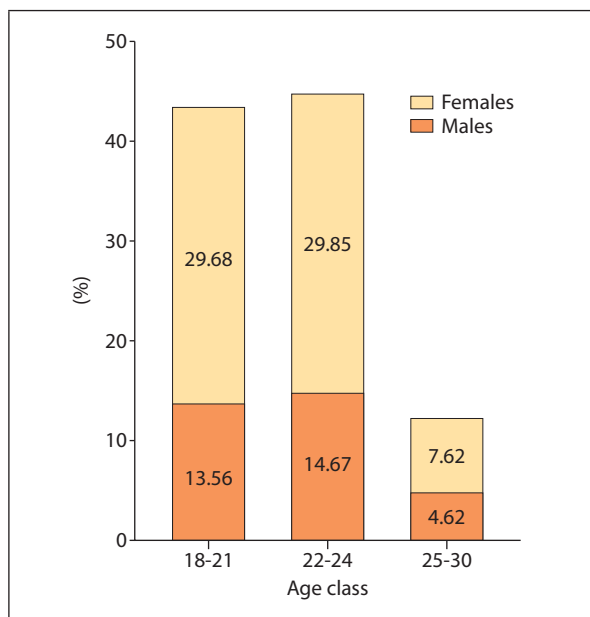


Figure 1
Combined distribution of responders by age class and sex (8325).

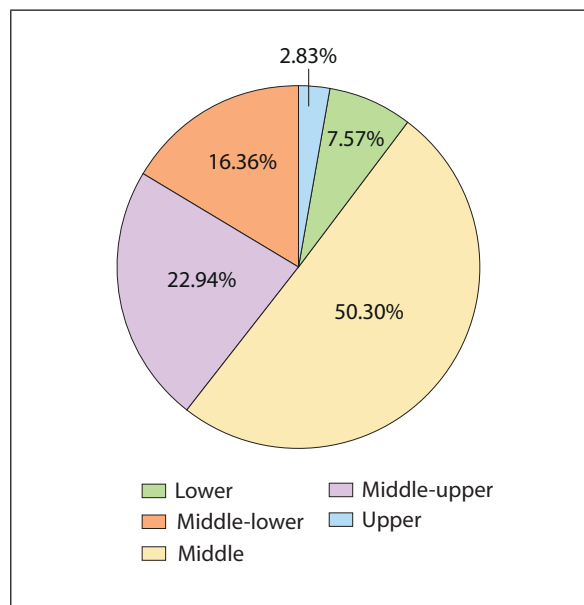


Figure 2
Distribution of responders by socio-economic status (8410).

Table 2

Combined distribution of responders by age class, sex and socio-economic status

| Socio-economic status | Age class | | | | | |
|-----------------------|-------------|-----------|-------------|-----------|-------------|-----------|
| | 18-21 | | 22-24 | | 25-30 | |
| | Females (%) | Males (%) | Females (%) | Males (%) | Females (%) | Males (%) |
| Lower | 1.7 | 0.6 | 2.5 | 0.9 | 1.3 | 0.6 |
| Middle-Lower | 4.5 | 1.6 | 5.9 | 2.1 | 1.5 | 0.9 |
| Middle | 15.7 | 7.0 | 14.9 | 7.1 | 3.7 | 2.0 |
| Middle-Upper | 6.9 | 3.8 | 6.0 | 4.1 | 1.1 | 1.0 |
| Upper | 0.7 | 0.7 | 0.6 | 0.5 | 0.1 | 0.2 |

the squared height (m). Furthermore, both age and BMI were coded in classes. For age, three classes were chosen: 18-21; 22-24; > 24. With respect to BMI, the WHO classification was adopted (< 18.5: underweight; 18.5-24.9: normal; 25.0-29.9: overweight; ≥ 30: obese) [3].

In order to identify incomplete, incorrect, inaccurate and irrelevant data, data cleaning was done by two researchers together. First, data of responders younger than 18 years or older than 30 years (218 observations) were removed. Furthermore, incorrect information about weight (< 10 kg or > 200 kg), height (< 120 cm or > 220 cm) and BMI (> 50) were coded as missing.

The final sample was composed by 8516 observations. The mean age of responders was 22.2 (Standard Deviation 2.0) and 5702 (67.0%) were females. Data were missing for 190 and 3 students respectively for age and sex. According to the distribution in age classes, 3601 (43.3%) belonged to the 18-21 one, 3796 (44.5%) to the 22-24 class and 1019 (12.2%) to the 25-30 class.

Figure 1 shows the combined distribution by age class and sex.

With respect to socio-economic status, data were available for 8410 responders and showed that 50.3% of students belonged to the middle class (Figure 2).

Table 2 reports the combined distribution of socio-economic status, age class and sex.

LIMITS AND STRENGTHS OF THE STUDY

The “Sportello Salute Giovani” project has several pitfalls [4]. First of all, people were voluntary enrolled in the survey. This may determine a selection bias because responders could have better attitudes towards health problems and more healthy behaviours [5]. As a consequence, some results about risk factors could be diluted. Furthermore, neither formal sample size calculation nor standardized sampling methods were used. Notwithstanding it, the final large sample is expected to guarantee the representativeness. The study may be also affected by misclassification, both non differential and differential [6, 7]. In fact, studies performed by means of questionnaire may lead to a recall bias. However, since this monograph looks at descriptive data, misclassification may be ignored.

As strengths are concerned, this study is the first to address, at the same time, several issues, in particular attitudes and behaviours toward health, in Italian university students with a so large sample size and a multi-center approach.

Conflict of interest statement

None.

Submitted on invitation.

Accepted on 16 April 2015.

REFERENCES

1. Università Cattolica del Sacro Cuore. “Sportello salute giovani”. Available from: http://brescia.unicatt.it/SportelloSaluteGiovani_BS.pdf.
2. Cavallo F, Giacchi M, Vieno A, et al. Studio HBSC in Italia (Health behavior in school-aged children): rapporto sui dati 2010. Roma: Istituto Superiore di Sanità; 2013. (Rapporti ISTISAN, 13/5). Available from: www.iss.it/publ/index.php?lang=1&id=2685&tipo=5.
3. World Health Organization (WHO). BMI Classification. Available from: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html. Geneva: WHO.
4. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002;359(9302):248-52.
5. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*. Philadelphia: Lippincott Williams & Wilki; 2012.
6. Ljungvall Å, Gerdtham UG, Lindblad U. Misreporting and misclassification: implications for socioeconomic disparities in body-mass index and obesity. *Eur J Health Econ* 2015;16(1):5-20.
7. Rasmussen M, Holstein BE, Melkevik O, Damsgaard MT. Validity of self-reported height and weight among adolescents: the importance of reporting capability. *BMC Med Res Methodol* 2013;13:85.