Escola Nacional de Saúde Pública
UNIVERSIDADE NOVA DE LISBOA

Measurement error in self-reported risk factors for cardiovascular disease: results from the first Portuguese National Health Examination Survey

Doctoral Programme in Public Health

Specialization in Epidemiology

Irina Kislaya

October 2021

Escola Nacional de Saúde Pública
UNIVERSIDADE NOVA DE LISBOA

Measurement error in self-reported risk factors for cardiovascular disease: results from the first Portuguese National Health Examination Survey

Dissertation submitted in compliance with the requirements for the degree of Doctor of Public Health - Specialization in Epidemiology, carried out under the scientific supervision of Professor Baltazar Nunes and Professor Hanna Tolonen

October 2021

Abstract

Accurate data on cardiovascular disease risk factors are essential for the design and evaluation of public health policies. Currently, self-reported data on hypertension and hypercholesterolemia constitute a primary data source for population health monitoring at European and national level.

This thesis aimed to evaluate accuracy of self-reported data on hypertension and hypercholesterolemia, to quantify, and to correct measurement error bias using data from the first Portuguese Health Examination Survey (INSEF).

Measurements of blood pressure and total cholesterol concentrations were used as gold-standards to estimate measurement error bias, sensitivity and specificity of self-reports, investigate error impact on outcome-exposure associations and illustrate application of multiple imputation for bias correction.

Despite all the efforts to limit the error through design and fieldwork procedures, self-reported data underestimated prevalence of hypertension and hypercholesterolemia in Portuguese population. Being unequally distributed among socioeconomic subgroups, measurement errors resulted in underestimation of socioeconomic inequalities in younger and overestimation in older age groups. The study results highlight the importance of measurement error bias analysis when using self-reported data. Results from multiple imputation show the approach feasibility for measurement error bias adjustment in prevalence estimates and outcome-exposure associations when individual-level validation data is available.

Health statistics on cardiovascular disease risk factors derived from self-reports should be used with caution. Integration of objective measurements in large-scale health surveys will improve the accuracy of epidemiological information on hypertension and hypercholesterolemia.

Keywords: measurement error, misclassification bias, self-report, hypertension hypercholesterolemia

Resumo

Informação epidemiológica de qualidade sobre os fatores de risco para doenças cardiovasculares é essencial para formulação e avaliação das políticas de saúde. Atualmente, os dados autoreportados sobre hipertensão e hipercolesterolemia constituem uma fonte primária de informação para monitorização da saúde da população a nível europeu e nacional.

O objetivo desta tese é avaliar a qualidade dos dados autoreportados sobre hipertensão e hipercolesterolemia, quantificar e corrigir o viés de medição assocados à informação autoreportada utilizando os dados do primeiro Inquérito Nacional de Saúde com Exame Físico (INSEF).

Considerando medições diretas da tensão arterial e do colesterol total como padrão, estimou-se a sensibilidade e a especificidade dos dados autoreportados, avaliou-se o impacto do erro da medição nas estimativas da prevalências e medidas de associação. Adicionalmente, o trabalho desenvolvido ilustrou a aplicação de imputação múltipla para correção de viés de medição.

Os dados autoreportados subestimaram a prevalência de hipertensão e hipercolesterolemia na população portuguesa comparativamente às medições diretas. O grau diferencial de viés por estatuto socioeconómico leva a subestimação das desigualdades socioeconómicas nos mais jovens e superestimação nos grupos com a idade mais avançada.

Os resultados do estudo realçam a importância da análise de viés associados à medição na utilização dos dados autoreportados. Os resultados da imputação múltipla indicam a viabilidade desta estratégia no ajuste das estimativas da prevalência e medidas de associação quando estão disponíveis a nível individual os dados das medições diretas.

As estatísticas de saúde derivadas dos dados autoreportados devem ser usadas com cautela. A integração das medições objetivas nos inquéritos de saúde de base populacional permitirá obter informação epidemiológica de melhor qualidade para monitorização da hipertensão e hipercolesterolemia.

Palavras-chave: erro de medição, viés de classificação, autoreport, hipertensão arterial, hipercholesterolemia

Acknowledgments

I would like to sincerely thank my supervisors Prof. Baltazar Nunes and Prof. Hanna Tolonen for their time, guidance and support throughout this project.

I am also very grateful for the advice and encouragement I received from my other advisory committee member, Prof. Julian Perelman.

I would like to thank National Health Institute Dr Ricardo Jorge for sharing the data used in my PhD. Special thanks to Prof. Carlos Dias, the Head of the Department the Epidemiology of National Health Institute Dr Ricardo Jorge who motivated me to pursue PhD.

A special thanks to Andreia Leite, Ausenda Machado, Ana João Santos and João Santos who has helped with the kind and generous advice over the years.

I acknowledge and am thankful for the contribution of the co-authors of the papers in this thesis.

Finally, I would like to express my deepest gratitude to my friends, to my family, especially to my husband Dmitry, for all their love, patience, faith and support throughout this long journey.

Table of contents

List of Tables

List of figures

List of Abbreviations

AUC - area under receiver operating characteristic curve

EU - European Union

ECHI - European Core Health Indicators

HES – health examination survey

HIS – health interview survey

GP - general practitioner

FN – probability of false-negative

FP - probability of a false-positive

IC95 - 95% confidence interval

INS - National Health Interview Survey

INSEF – Portuguese National Health Examination Survey

ISCED 2011 - 2011 International Standard Classification of Education

KPI – Key performance indicators

MIME - multiple imputation for misclassification error

MSE - mean squared error

NHANES - The USA National Health and Nutrition Examination Survey

NHIS – The USA National Health Interview Survey

NUTs – Nomenclature of territorial units

OECD - Organisation for Economic Co-operation and Development

pp - percentage points

PR - prevalence ratios

RII – relative inequality index

SE - standard error

SII - slope index of inequality

SP - specificity

TSE - Total survey error

1. Introduction

The measurement of populations' health is keen to address overall health improvement, reduce inequalities and strength healthcare systems (1–4). However, gathering high quality health-related data to inform policies remains a challenging task, particularly at population level. To provide insight on population health and address global health challenges health information systems make a use of multiple data sources. Among those, even in an era of Big Data, population-based health Interview surveys (HIS) that collect self-reported data using questionnaires, continue to play an important role (2,4).

HIS are widely used for public health surveillance and cross-country comparisons at the global, regional and national level. For instance, in the European Union (EU) about 25% of policy-focused standardized indicators for health monitoring (European Core Health Indicators - ECHI) are based on self-reported data from HIS (2,4). In Portugal, national HIS, denominated by the National Health Interview Survey (*Inquérito Nacional de Saúde* - INS), for decades represents an important tool for the monitoring and evaluation of National Health Plans (5) and Priority Health Programs. INS was established in early 1980s and has been continuously developed since when (6). Currently, INS integrates the European Health Statistics System as a Portuguese component of European Health Interview Survey (7,8).

HIS can be implemented in large-scale using probabilistic samples of population and combines information on disease status, health determinants, health believes, access to health care and health expenditures with socioeconomic characteristic, rarely jointly available in other data sources within modern health information systems. These feathers are the core strengths of HIS.

In the past decades, the demand for high-quality health survey data has increased, including aspects such as estimates precision, levels of desegregation, accuracy and sustainability. Particular attention has been paid to methodological issues related to survey error (9). It is widely acknowledged in literature that self-report as a method of data collection presents some weaknesses (10–15), it can be subject of measurement error. Measurement error in survey arises when registered participant's response for any reason is different from true respondent status. In epidemiology and survey research, it is recognized as one of the most damaging sources of error that can increase either the uncertainty of estimates or bias estimates from true population values, with serious implications for quality of research and decision-making.

The reporting error in HIS can be particularly relevant for the measures of existence of the clinical conditions and health determinants (10–15). Recall problems, undiagnosed disease or misunderstanding of medical terms and social desirability are recognized as the major sources of measurement error associated with self-reports (10,16). As such, statistical inference on prevalence of disease and health determinants based on self-reports from HIS may not truly represent what is happening in the population. In addition, some population subgroups can be more susceptible to reporting errors than the others. In such cases, statistical inference from HIS will also result in underestimation or overestimation of true association between health outcomes and their social determinants. This can be particularly relevant for research on health inequalities (17,18), important part of nowadays public health agenda (19). In Portugal little is known about measurement error associated to self-reports.

In order to obtain accurate and reliable information on population health through HIS it is essential to address four measurement error dimensions (Figure1):

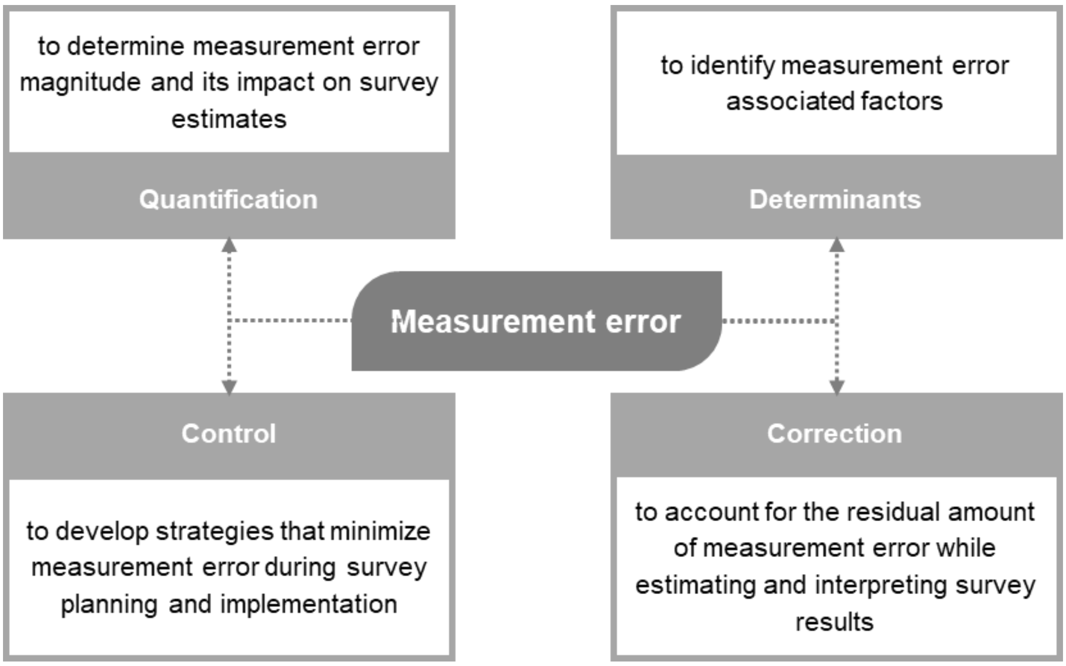


FIGURE 1 MEASUREMENT ERROR DIMENSIONS TO BE ADDRESSED

Health Examination Surveys (HES) that collect both self-reported and objectively measured data (biomarkers) may provide more accurate information than HIS, relying exclusively on self-reported information (10,20–22). HES present an opportunity to perform research on the accuracy of self-reports and can provide better understanding of measurement error in estimates, improving the results interpretation important for credibility of the health surveys.

In last decades, HES has become an important part of the health information systems worldwide. Between 2007 and 2021 in the EU almost half of member states have gained an experience of conducting HES at national or regional level (21,23). In Portugal the first national HES (the Portuguese National Health Examination Survey (Inquérito Nacional de Saúde com o Exame Físico - INSEF) was developed in 2015 (24). It provides valuable combination of self-reported and objectively measured data obtained through the component of physical examination and blood collection for the same participants, making possible to gain insight on degree of measurement error associated to self-reports in Portuguese population.

The studies presented in this thesis intent to contribute for an improvement of measurement of health indicators in Portugal by focusing on component of total survey error associated to data collection methods in HIS surveys, i.e., the measurement error. They make use of data from two national surveys developed in 2014-2015: The first is INSEF that combines self-reported and objectively measured data, and the second is a national HIS (INS2014) that relies solely self-reports. As a case study, I considered two modifiable cardiovascular disease risk factors for which both objectively measured and self-reported data was available: high blood pressure (hypertension) and elevated cholesterol (hypercholesterolemia).

This thesis is organized in seven chapters, starting with introduction and background that provide overview on measurement error in epidemiology and survey research introducing key concepts, motivation for selection of self-reported health indicators for research and present summary overview of existing literature on validity of self-reported data. Chapter 3 contains research objectives. Chapter 4 describes the methods used in this thesis. Chapter 5 presents the results. This thesis used research-paper model and the results of developed studies were published as following research papers in peer-reviewed public health journals:

- **Paper 1:** Kislaya, I., Tolonen, H., Rodrigues, A.P., Barreto, M., Gil, A.P., Gaio, V., Namorado, S., Santos, A.J., Dias, C.M., Nunes, B. (2019) Differential self-report error by socioeconomic status in hypertension and hypercholesterolemia: INSEF 2015 study. European Journal of Public Health, Vol. 29, No. 2, 273–278, https://doi.org/10.1093/eurpub/cky228
- **Paper 2:** Kislaya, I., Perelman, J., Tolonen, H., Nunes, B. (2019) Do self-reported data accurately measure health inequalities in risk factors for cardiovascular disease? International Journal of Public Health, 64(5), 721-729, doi:10.1007/s00038-019-01232-1

- **Paper 3:** Kislaya, I., Santos, A.J., Lysol, H., Antunes, L., Barreto, M., Gaio, V., Gil, A.P., Namorado, S., Dias, C.M., Tolonen, H., Nunes, B. (2020) Collecting valid and reliable data: fieldwork monitoring strategies in a Health Examination Survey. Portuguese Journal of Public Health. 2020;38:81–90. DOI: 10.1159/000511576
- **Paper 4:** Kislaya, I., Leite, A., Perelman, J., Machado, A., Torres, A. R., Tolonen, H., Nunes B. (2021) Combining self-reported and objectively measured survey data to improve hypertension prevalence estimates: Portuguese experience. Arch Public Health. 2021 Apr 8;79(1):45. doi: 10.1186/s13690-021-00562-y.

Chapter 6 provides an overall discussion of the findings, strengths and limitations and presents the implications of the further research and practice. Finally, the summary conclusions are presented in Chapter 7.

The studies presented in this thesis will provide a comprehensive view on measurement error in HIS addressing its outlined dimensions: Quantification, Determinants, Control, and Correction. By addressing these four dimensions of the measurement error in self-reported data this thesis contributes to the improvement of health information available for monitoring of hypertension and hypercholesterolemia in Portugal and, more widely, to the design and implementation of better-informed public health programs.

2. Background

In this Chapter, I present the theoretical background on measurement error starting from definition of measurement error in an epidemiological and survey perspectives. I describe error impacts on estimates and present an overview of approaches to its assessment, strategies to minimize it during studies planning and implementation, as well as and several statistical methods that can be employed to mitigate its effects. Further, I present motivation for selection of hypertension and hypercholesterolemia indicators and brief literature overview on measurement error in hypertension and hypercholesterolemia collected by surveys, identifying exciting research gaps.

2.1. Measurement error and misclassification in epidemiology and survey research

2.1.1 Definition

Surveys that gather self-reported data directly from respondents are widely used in observational epidemiology and public health surveillance. For epidemiologists and statisticians precision, validity, and generalizability of estimates are the key elements of survey inference since without it data have little or no use. However, no study is perfect and accurate estimates (valid and precise) are difficult to achieve. Errors can arise in different steps of study implementation and their presence can strongly affect estimates and compromise the validity of study conclusions.

In a broad sense, an error can be defined as a misleading or false result or any deviation of the study results from the truth regardless of its reasons (25).In epidemiology, errors are traditionally classified at random and systematic (25–28). This approach is applied to all data gathering methods and is not exclusive to surveys. A random error is the portion of the variation in an estimate that occurs due to chance and has no apparent connection to any collected variable or study procedure (26). It occurs mainly due to sampling variability and can be easily quantified and, to some point, controlled with study sample size, although sampling is not a unique source of random error in estimates. The amount of random error in estimates is usually described as estimates “precision” or “uncertainty” (26) and quantified through the confidence intervals estimation.

Systematic error, on the other hand, represents a tendency in a study design, measurement, or any other procedure leading to estimates that deviate from the true values in some particular direction (25). Systematic error in epidemiology is commonly referred to as "bias" (27) and estimates with a small amount of systematic error referred to as valid or unbiased. Three major contributors to systematic error in epidemiological studies are distinguished: selection bias, confounding, and measurement error (26,27).

Survey researchers also recognize that error can be random and systematic in nature(29), but use an alternative taxonomy to address it, linking the error to different steps of survey design and implementation. An excellent summary and historical evolution of the concept of survey error has been provided by Biemer et al (30).

Within the Total survey error (TSE) framework two major error groups are distinguished: errors affecting the measurement and the representativeness (Figure 2) each incorporating random and systematic components (9). The first group frequently referred to as measurement or observational error, encompasses errors related to the stages of the measurement process, including the design of the measurement instrument, its application to measurements, and registry and data processing. The second group, also denominated as non-observational error, encompasses errors related to the sampling frame coverage, sample selection, and non-response.

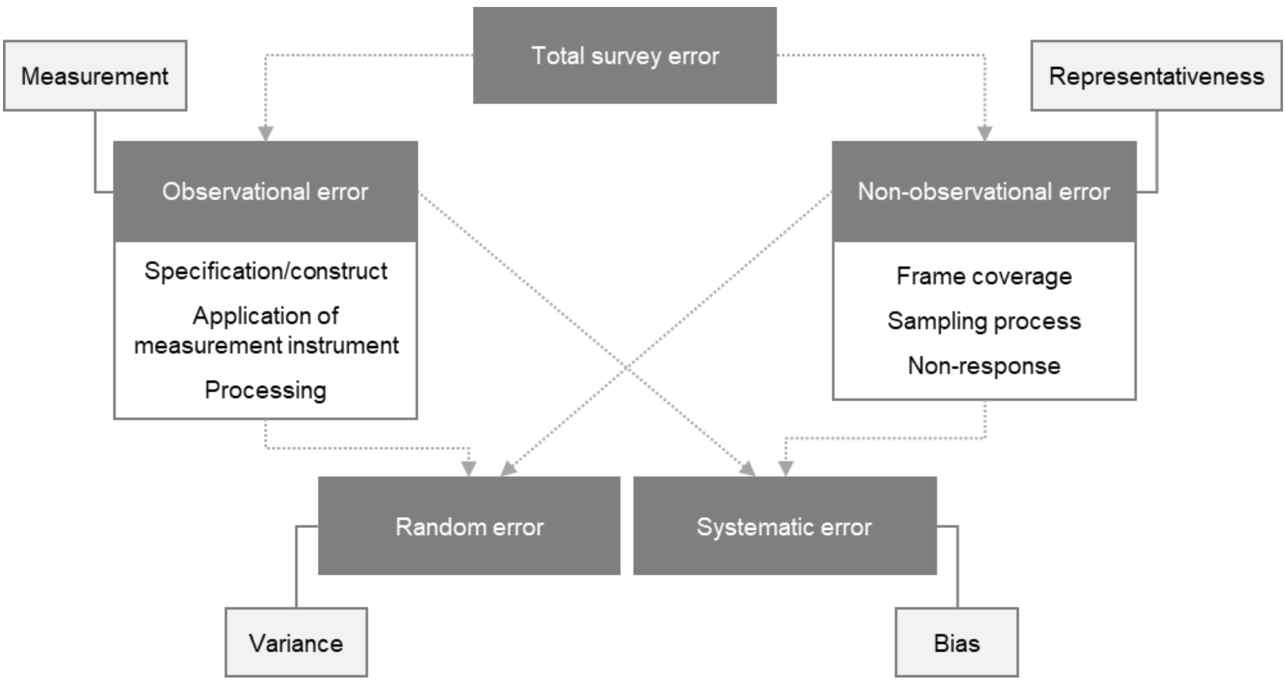


FIGURE 2 TOTAL SURVEY ERROR COMPONENTS (ADAPTED FROM GROVES ET AL)

Measurement error occurs during the data gathering process when the recoded value of the variable being measured is incorrect and diverges from the true value. In a survey, it can arise for a variety of reasons, including i) deficient survey question formulation or indicator operational definition, ii) deficiencies during the data gathering process, such as improper interviewing or inadequate use of the measurement instrument, iii) incorrect registry or data coding, iv) misreporting or untruthful answer to survey question by the respondent (30,31). Although measurement error can be both random and systematic contributing to extra uncertainty and bias in estimates, its systematic component is more substantial than the random (26,29).

In survey research, the term measurement error applies to variables regardless of the measurement scale, in epidemiology when a measured variable is categorical (nominal or ordinal), traditionally measurement error is denominated as misclassification (26,27). Misclassification can be defined as a specific type of measurement error in a categorical variable that occurs then study subjects are erroneously assigned to a different category than they should belong (27). For instance, cases classified as controls or subjects suffering from the disease considered healthy, or exposed subjects considered unexposed and vice versa are classical examples of misclassification in epidemiological research. The focus of this thesis will be a particular case of misclassification in binary variables.

2.1.2 Approaches to assessment

The accuracy of estimates can be described in terms of the mean square error (MSE) metric. MSE is defined as the expected squared difference between an estimated statistic and the population parameter it is intended to estimate, which can be decomposed into the sum of the squared bias and the variance (9,31):

$$MSE = Bias^2 + Variance$$

corresponding to systematic and random error components, respectively.

This metric reflects the cumulative contribution of all error sources, including measurement errors, large MSE indicates that one or more sources of error have an adverse effect on the estimate, while small MSE indicates that the estimate is accurate.

Variance component associated with the measurement process can be estimated through repeated application of the measurement instrument, while for assessment of bias collection of additional data is required (28,29,32). Traditionally, systematic error associated with the measurement process is estimated comparing an imperfect, error-prone measurement to the reference or gold-standard method, i.e. the best available conventional clinical assessment, which is assumed more accurate or error-free (28–30,32). When addressing the accuracy of self-reported survey data, researchers often consider medical records, administrative registries, objectively measured data and biomarkers as gold-standard (30,33–35). Such data is obtained from specifically designed validation studies that collect information on a variable of interest sing error-prone and reference measurement methods alongside the same set of study subjects allowing researchers to understand and quantify the impact of mismeasurement, establishing the relationship between error-prone and gold-standard measurements. This relationship is designated by a measurement error model (28,32,36) and, for categorical variables, can be specified by classification probabilities (32,37).

Let us consider a binary outcome variable Y (e.g., having a disease /not having a disease) that has been measured perfectly without error (true outcome, reference measure) and its error-prone version Y* (self-reported outcome) available for the same study participants from a survey subject to misclassification (Table 1). Based on the reference measure and self-assessment results, participants can be assigned to one of the four cells as shown below:

TABLE 1 EXEMPLIFICATION OF CROSS-CLASSIFICATION OF TWO MEASUREMENTS

	True status (gold-standard measure)		
Result of classification self-report		Y=1	Y=0
	Y*=1	a	b
	Y*=0	c	d

Performance of error-prone measurement method can be assessed through estimation of classification probabilities based on constructed two-by-two contingency table.

The probability of being correctly classified having a condition, given that participant is known to have the condition according to the gold-standard measure, is denominated as sensitivity and defined (32) as:

$$\textbf{Sensitivity}\ (\textbf{SE})\text{: } \Pr\,(Y^{*}=1|\mathrm{Y}=1)=\frac{a}{a+c}$$

The probability of being correctly classified as not having a condition given that participant does not have the condition according to the gold-standard measure is similarly defined as:

$$\textbf{Specificity (SP)}: \Pr(Y^* = 0|\mathrm{Y} = 0) = \frac{d}{d+b}$$

The probability of incorrect classification is denominated as the probability of a false-negative (FN) and the probability of a false-positive (FP) can be obtained directly from values presented in the contingency table or derived from estimates of sensitivity and specificity:

$$\textbf{FN} = \Pr(Y^* = 0|\mathrm{Y} = 1) = \frac{c}{a+c} = 1 - SE$$

$$\textbf{FP} = \Pr(Y^* = 1|\mathrm{Y} = 0) = \frac{b}{d+b} = 1 - SP$$

In research on the accuracy of self-reported data FN and FP are frequently of main interest for researchers and traditionally referred to as underreporting and overreporting errors, respectively (32)

Another important statistic also used in validation studies, although less frequently, are reclassification probabilities or predictive values (28,32) The positive predictive value (PPV) is the probability that a true outcome occurred, given that the participant reported the outcome to occur. The negative predicted value is defined as the probability that a participant actually did not experience the outcome given that the participant reported not to experienced it.

$$\textbf{PPV} = \Pr(\mathrm{Y} = 1|Y^* = 1) = \frac{a}{a+b}$$

$$\textbf{NPV} = \Pr(\mathrm{Y} = 0|Y^* = 0) = \frac{d}{c+d}$$

These quantities as well as sensitivity and specificity can be used to perform bias assessment and algebraic correction of estimates of outcome distribution in the population or outcome-exposure association measures (32,38).

2.1.3 Impact on estimates

It is commonly acknowledged that measurement error and misclassification can strongly affect parameters estimates. Its consequences are widely covered in the literature, an excellent overview on the topic was presented by Lash (28), Buonaccorsi (32), Greenland (39), and Keogh (37). Most research, however, focuses on the impact of misclassification in exposure (37,40), misclassification in outcome variables has received less attention in the literature (26), although some studies have discussed and proposed approaches to correcting it (34,37,41–48).

It has been shown that outcome misclassification has the potential of underestimating or overestimating prevalence and incidence as well as the measures of association (odds ratio, risk ratio, etc). The direction and magnitude of bias caused by imperfect measurement of the outcome depend on the misclassification mechanism and misclassification probabilities (26,38).

Commonly, mismeasurement originates both false positive and false negative. For prevalence and incidence estimates, when false-negatives occur more frequently than false-positives, the frequency of the outcome in the population will be underestimated. Conversely, a higher number of false-positive compared to a false-negative yields overestimation.

For outcome-exposure associations, the effect of outcome mismeasurement is not straightforward and when addressing it, it is important to distinguish between differential and non-differential types of error (27,37,38,49). Differential misclassification occurs when some study subjects are more prone to being misclassified than others or, in other words, when the degree or direction of error differs between study groups concerning any other variable being measured In such cases, sensitivity and specificity and false-positive and false-negative rates are not uniform and vary among studied subgroups. When the probability of individuals being misclassified does not depend on other variables, being the same among all studied population subgroups, the misclassification mechanism is described as non-differential.

In a presence of non-differential outcome misclassification bias in outcome-exposure associations generally are expected to be towards the null (27,50). Assuming a non-differential misclassification mechanism, it is expected to obtain a parameter estimate that points out in the same direction as an association measure based on a perfect measurement method but of smaller magnitude, as indicating by a reduced strength of association. Researchers put considerable effort during the study design and implementation stage to guarantee this assumption. However, even if the misclassification mechanism is non-differential, the bias is not always towards the null (36,49,50). In fact, it has been demonstrated the association estimate can even be reversed, when the amount misclassification is severe the exposure is not binary or the exposure or other covariates are also measured with error (49).

With differential outcome misclassification, the exposure will affect the error-prone measurement of Y* through different pathways than the true outcome Y. In this case, bias in outcome-exposure association estimates introduced by imperfect measurement, can be in any direction reducing or increasing the apparent strength of association (26) and careful bias assessment, on a case by case basis, is required.

2.1.4 Determinants or sources of error

Identification of factors that can cause measurement error in survey is a first step to minimize it and reduce its effect on the estimates. Factors affecting measurement process in survey have been grouped into four domains: the questionnaire features, the data collection method, the interviewer, and the respondent characteristics (Figure 3) (29).

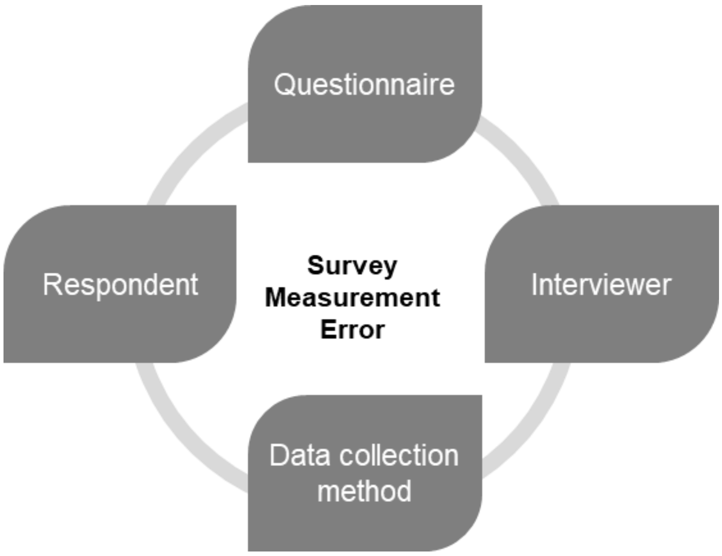


FIGURE 3 SOURCES OF MEASUREMENT ERROR IN SURVEY

Questionnaire design is an important step of the survey process that has a strong influence on validity of the inference. A language, wording of the questions, questions type (close-ended or open-ended), available response options, order of individual questions, skip rules, as well as the questionnaire visual layout and general questionnaire length can affect accuracy of the information obtained by the questions as well as overall participants´ willingness to respond (31,51–53).

Several features of the data collection method are also recognized in the literature as important contributors to measurement error in surveys (30,31,54–57). Those include the administration mode (self-administrated without the aid of an interviewer, by phone or by face-to-face interview), the format (in-paper or electronic), devices used for the administration (computer, tablet, smartphone) as well as the interview context (at home, at workplace, in a clinic, and other) (30,31,54–57).

Survey measurement errors also can derive from the interviewers' characteristics and behaviors (58–61). In all interviewer-mediated surveys, interviewers play a crucial role, given they make contact with and gain respondents' cooperation, ask questions, conduct measurements, record answers and measures, and maintain respondents' motivation throughout the interview (58–60). An interviewer can introduce error by adopting inadequate interviewing styles, probing for participant response, or making judgment errors when recoding participants' responses. Interviewers' characteristics (sex, age, tone of voice, etc.) and even their general appearance can also influence participants' cooperation and quality of response (58,59,62).

Likewise, respondents' characteristics also play an important role in a response process (35,63,64). Several studies have demonstrated that respondents' age, sex, education, income, employment, personal experiences, knowledge and attitudes, culture can influence questions interpretation and affect the quality of reporting (14,15,35,64–69).

2.1.5 Control and Correction: summary of selected methods

Standard research practices aimed to avoid and minimize errors due to imperfect measurements during the survey design and implementation. These practices include careful design and evaluation of survey questions by a combination of qualitative and quantitative methodologies, including expert review, focus groups, cognitive interviews, eye-tracking, interviewer debriefing, pretest, and piloting. An excellent summary of a current state of the art on the topic, an overview of software solutions available to assist researchers on these tasks, and guidance on when and which of these methods should be used has been presented by Robinson & Firth (53) and Beatty et al (52).

Another important step towards measurement error reduction is a standardization of all survey procedures (22,31,70). This can be achieved using: i) a manual of survey field operations; ii) adequate interviewers training; and iii) quality assurance and control measures. The manual of survey field operations should include a comprehensive description of all survey procedures, that will guide interviewers and other survey staff during the data collection and processing helping them to address arising problems and questions in a standardized consistent manner. Proper interviewer training encompasses demonstrations, role-playing, and piloting. This allows the interviewer to administrate the questionnaire in real-life conditions but with close supervision of senior staff (58). Another important aspect is the implementation of quality assurance and control measures during the course of fieldwork (31,62). For interviewer-mediated surveys, fieldwork supervision traditionally includes activities of direct interviewer observation and systematic evaluation of adherence to the standard survey procedure through collection and analysis of paradata. Several key performance indicators (KPI) (71) have been proposed in the literature to monitor interviewer performance, including, individual workload, contact and cooperation rates, interviews duration patterns among others. In addition, to evaluate adherence to standard interviewing procedure and detect interviewer misconduct, quality control systems traditionally include validation interviews, when fieldwork supervised re-contact selected respondents asking them for feedback on survey participation and re-administrating part of the questionnaire (31).

Identification of deviations from the standard procedures requires correction actions, that include individual debriefing sessions, focus groups discussions, and re-training.

Even with all these efforts, some residual amount of error due to mismeasurement will persist and need to be addressed after the data collection. To account for measurement error and misclassification, quantitative bias analysis, regression calibration, maximum likelihood-based methods, and multiple imputation have been proposed, among others (28,38,42–44,48,72).

Despite extensive theoretical knowledge and existing methodologies to identify and account for, researchers rarely pay proper attention to systematic errors originated by imperfect measurements (37,40). Traditionally, a simple disclaimer in the study limitation section referring that "study is based on self-reported data and possess associated limitations" or general qualitative assessment such as "use of self-reported data may lead to underestimation of the effect" are considered sufficient by peers to get results published. The review (40) of original research published in 2016 in 12 high-impact medical journals showed that only 3% of studies investigated and quantified the possible impact of the measurement error during the data analysis.

Quantitative bias analysis

The most simple approach to estimate misclassification bias is algebraic that relies on previously defined misclassification probabilities. This method has been referred to in the literature as quantitative bias analysis (28).

Given values of sensitivity (SE) and specificity (SP), the estimate of proportion based on error-prone measurement ($\hat{p}$) can be corrected for misclassification as following (32):

$$\hat{\pi} = \frac{\hat{p}+SP-1}{SE+SP-1},$$

and bias can be determined as $bias = \hat{p} - \hat{\pi}$.

The algebraic approach can be also used to account for misclassification bias in measures of association estimated from two-by-two tables (28). To do so, cell frequencies are recalculated based on estimates of sensitivity, specificity and observed counts, and new adjusted cell frequencies are used to compute estimates of the measures of association (odds ratio, risk ratio, risk difference, etc.) as usual. Alternatively to sensitivity and specificity, proposed correction methodology also can rely on positive and negative predictive values estimates (28). Illustration of this approach has been provided in several textbooks (28,32) and papers (72) and implemented in Excel and most of the standard software packages for data analysis, such as R, Stata, SAS (28).

Estimation of 95% confidence interval for corrected measures of association based on adjusted cell frequencies has been proposed by Greenland et al (72). However, suggested methodology is based on random samples assumption and not account for sampling weights so frequently used in the analysis of large-scale population surveys due to complex sample design. Nevertheless, the application of correction formulas to point estimates in complex surveys settings may still be informative for researchers.

The main advantages of this method include the simplicity of application and available software solutions and its application on aggregated data levels with sensitivity and specificity values estimated either from external or internal validation studies. Furthermore, this method can be implemented even if real data on the misclassification mechanism is not available. Using quantitative bias analysis for a range of sensitivity and specificity values a researcher may make a judgment on the robustness of study findings to several amounts of misclassification error.

As the most relevant limitations, it can not be implemented when researchers are interested in confounder-adjusted estimates of measures of association and is not suitable when the exposure variable is continuous. Although it does not account for uncertainty around sensitivity and specificity estimates, the method can be easily implemented when researched are interested in crude estimates of measures of associations or disease prevalence.

Regression Calibration

In the 1990s regression-based methods that apply to individual-level data and enable to investigate more complex relationships between variables have been proposed to account for outcome misclassification. A regression calibration has been developed initially for classical measurement error correction in continuous variables and later extended to deal with misclassification. The idea of the method is to replace error-prone observed values of a variable with predicted values from the calibration model. Application of regression calibration requires validation subsample with error-prone and gold-standard measures available. Although this approach is most frequently used to deal with error in exposure in nutritional epidemiology, it has been successfully applied in several studies for correction of prevalence estimates of chronic conditions in surveys (34,73). It can be easily implemented in the context of complex samples surveys, as most statistical packages have routines to account for different selection probabilities in regression (74,75).

Likelihood-based methods

More recently, maximum likelihood methods have been developed to account for outcome misclassification in effect estimates. Magder&Hughes (76) proposed to modify the likelihood function incorporating information on the values of sensitivity and specificity when estimating parameters in a logistic regression model using an expectation-maximization (EM) algorithm. Lyles (48,77) extended the maximum likelihood approach for differential misclassification, allowing the incorporation of fixed values of sensitivity and specificity, as well as estimation of sensitivity and specificity based on internal validation data. Primary, the maximum likelihood methods were developed for logistic regression models, later Edwards et al (78). has demonstrated that the method can be successfully extended to other generalized linear models, such as Poisson regression (78). A modified maximum likelihood estimator has been also effectively employed by Tang et al. (47) in a more complex framework of joined misclassification of outcome and exposure variables.

Likelihood-based methods are very flexible and have shown good performance under a variety of settings (47,48,77,78). These methods are suitable for both differential and non-differential misclassification, allow both categorical and continuous exposures. Among primary barriers to implementation in practice, we should mention limited available software solutions. The application of likelihood-based methods requires advanced programming skills to directly specify the likelihood function that includes adjustments for sensitivity and specificity and is not straightforward for the majority of public health researchers and epidemiologists. Additionally, the method may be computationally intensive.

Multiple imputation

Multiple imputation methods proposed by Rubin (1987) for handling missing data have gained increasing attention for measurement error and misclassification correction (42–44). Assuming that true error-free measurements of outcome are not observed, measurement error can be handled within the missing data framework. Using the information on error-prone measurements and other covariates true missing outcome values can be imputed, further multiple created imputed data sets can be analyzed using standard statistical software and estimation results combined using Rubin's rules. To develop imputation model information on misclassification mechanism and relationship between error-prone and gold-standard measurement methods is required. It can be obtained either from a separate validation study or from an internal validation subsample (12,42,79,80). For a binary outcome, traditionally logistic regression is used as an imputation model (44). However, more sophisticated imputation techniques may also be employed to performed misclassification correction and simultaneously account for missing data (81,82).

As the main advantage of the method, we should mention its flexibility. Namely, multiple imputation is suitable to account for both differential and non-differential misclassification, as well as for measurement error in continuous variables, and can be used regardless of the variable role in the study (exposure, outcome, covariate). It also can be used to correct prevalence estimates, crude and adjusted association measures (odds ratios, risk ratio, hazard ratios and others). Recent methodological developments allow the implementation of MIME under complex survey settings (80,83). Currently, standard statistical software packages (Stata, R) allow sampling weight both in constriction of imputation models and analysis of multiply imputed data, as well as account for clustering. In addition, implementing the MIME correction is straightforward and it does not require special programming skills.

Several practical applications and simulation studies on MIME performance can be found in the literature. MIME in exposure was described by Cole (43) in a framework of Cox regression. Ni et al (44) illustrated MIME application to reduce outcome misclassification bias in hazard ratios. Edwards et al (42) applied MIME to successfully account for outcome misclassification bias in confounder-adjusted odds ratios and risk ratios based on internal validation samples. Schenker et al (12) used MIME for correction of prevalence estimates and logistic regression coefficients in large-scale health surveys relying on external validation samples.

Among MIME limitations, several authors mention its computational intensity, as a number of imputations increase (81) and suboptimal performance with a sample size of validation data (44).

2.2. Misclassification in survey data on self-reported hypertension and hypercholesterolemia

2.2.1 Motivation for indicators selection

Cardiovascular diseases (CVD) are still the major contributors to the burden of disease globally and across Europe. (84–87). In Portugal, CVD are also a major public health problem, accounting in 2015 for 32443 deaths (29.8% of all mortality) (88). In addition, these conditions have relevant socioeconomic burden. In specific, in 2015 about € 1,174,407 were spent on health care of people with CVDs in Portugal. Estimated indirect costs for national economy due to productivity losses and informal care of people with CVDs were € 476,927 and € 878,833 respectively (86).

Many CVDs can be prevented through interventions on health determinants and risk factors. High blood pressure and elevated blood cholesterol are well-known and documented risk factor for CVDs: atherosclerosis, heart disease, stroke and other circulation problems (86). The Global Burden of Disease project has estimated that in 2015 high blood pressure and high total cholesterol accounted for the loss of 211.8 and 88.7 million disability adjusted life years, respectively (84).

In Portugal CVDs are considered priority for public health action, National Priority Health Programs are dedicated to prevention and control of CVDs and its risk factors (89) with particular focus on hypertension and hypercholesterolemia (90,91). Timely diagnosis and effective management of these conditions are crucial for improvement in overall population health and reduction of premature mortality due to CVDs. At population level, accurate data on hypertension and hypercholesterolemia are essential for evidence-based public health planning, development and evaluation of national health problems.

In Portugal and across EU, hypertension and hypercholesterolemia are frequently measured by self-reports in epidemiological research and public health surveillance. Currently, self-reported data is a preferred data source for hypertension and hypercholesterolemia monitoring at national level, although several epidemiological studies had measured the prevalence of hypertension and hypercholesterolemia in Portugal objectively (92,93) .

At EU level, these indicators are monitored based on HIS. Prevalence of high blood pressure constitutes one of the harmonized European Core Health Indicators continuously monitored at the in EU since 2008 (2) using European Health Interview Survey data that provides high level of methodological standardization and comparability between countries. Self-reported indicator of lipid abnormalities is novel at EU level, it was introduced in the most recent wave of European HIS, developed in member states in 2018-2020 (94). Its current formulation make no distinction between different factions of lipid profile, defining hyperlipidemia as having high cholesterol, high blood lipids or triglycerides in the past 12 months. Evaluation of measurement error associated with self-repot for Portugal have particular relevance and can contribute to improvement of operational definition and data harmonization of indicators for lipid abnormalities monitoring in future surveys at EU level.

Both hypertension and hypercholesterolemia have a long asymptomatic period; these conditions are denominated as “silent killers” as most affected individuals experience no symptoms at all. These add extra complexity to timely diagnosis and correct reporting. Individuals may not be aware of their health status and, therefore, cannot report it accurately. If this is the case, inferences about the general population health based on self-reported data on hypertension and hypercholesterolemia can be incorrect with important implications for public health policy formulation, programs evaluation and cross-country comparisons.

For both, hypertension and hypercholesterolemia, there are biomarkers that can be used to measure these conditions objectively. Application of two measurement approaches to a sample of the same individuals will provide valuable information on degree of reporting error in survey data. Using hypertension and hypercholesterolemia as an illustrative example, this research will contribute to better comprehension of limitations and strengths of existing data sources to inform public health policies.

2.2.2 Measurement error in self-reported hypertension and hypercholesterolemia: a review of existing evidence

With the intensive technological development of biomarkers and decreasing costs of its collection on large scale, objective measurements of health conditions are integrated into surveys more frequently. HES are becoming an important part of the health information systems worldwide. Some countries such as Finland and USA have the longest experience of HES implementation (since 50s-60s) while others make their first steps on HES development (21,23,95). Although implemented less frequently and on a smaller scale, HES that in addition to interviews assemble objective measures of health through physical examination and collection of biological specimens are extensively used to perform studies on the accuracy of self-reported information (13,96–98). The most popular methodological approach consists in a direct comparison of a self-reported measurement with a biomedical measurement for the same condition, which is likely to be more accurate. This approach has been widely used to assess the accuracy of self-reported hypertension (11,13,15,67,96), diabetes (15,67,96), hypercholesterolemia (96,99), tobacco consumption (98) and other health behaviors (100). Other popular methods applied in accuracy studies include the comparison of self-reported data with administrative registries (101) and the confirmation of diagnosis with the general practitioner (102).

In this section, I will present a brief overview of existing research on accuracy of self-reported hypertension and hypercholesterolemia. Most of the studies were included in recent systematic reviews by Paalanen et al (33) and Golçalves et al (103). As such, in addition to the main research findings, I will focus on the methodological aspects of the studies that were not yet covered in order to evaluate studies comparability and identify existing research gaps on this topic. The following methodological aspects will be considered:

i) study design,
ii) target population;
iii) sampling design and sample size;
iv) measurement protocol and operational definitions of self-reported and objectively measured indicators, including cut-off points for objective measures;
v) statistical methods used to access self-report accuracy;
vi) whether study address factors associated with accurate or incorrect self-report;
vii) whether study address adjustment of estimates for measurement error.

Summary information on mentioned methodological aspects is presented in Appendix A.

Further along this thesis direct measurement of blood pressure and total cholesterol concentration will be referred to as objectively measured or examination-based, these terms will be used as synonyms.

All identified studies used a cross-sectional design focusing on the assessment of self-report accuracy in comparison to the examination-based data in a single time point.

Regarding target populations, studies present a variety of settings regarding participants´ age. In particular, for hypertension, almost half of validation studies focused on individuals of older age (45+ or 50+), others considered individuals aged 18 years and older or established different lower and upper bounds of age for the target population. Similar heterogeneity of age inclusion criteria was also verified in hypercholesterolemia studies. Additionally, some of the studies were limited to a particular population subgroup, such as, female health professionals (104) or individuals at higher risk of cardiovascular disease (105,106).

Sample sizes in the considered research also varied notably ranging from 216 to 13610 for hypertension and from 439 to 24 069 for hypercholesterolemia. Most of the studies were conducted with random samples or even using a complex sampling design, even though some (67,106) were developed using non-probabilistic samples.

A review of existing studies revealed a lack of nationally representative results across Europe for both hypertension and hypercholesterolemia. The majority of available data on the accuracy of hypertension was provided by regional (15,107) or local (10,67,106,108) studies, with no representativeness at the national level. National level data are available for Ireland (109,110), but only for older individuals. For hypercholesterolemia, nationally representative samples were used to evaluate self-reports accuracy in Korea, the USA, Ireland and Australia (68,96,99,111).

In Portugal, several studies used objective measurements of hypertension and hypercholesterolemia (92,93,112,113). However, these studies focus mainly on estimation of prevalence, some also estimated awareness of the disease (that is equivalent to sensitivity defined in a section 2.1.2). To our best knowledge, only one local study (10) with a small sample size has been published with Portuguese data on measurement error in surveys. A broad assessment of measurement error in self-reported hypertension and hypercholesterolemia has not been performed yet for the Portuguese population. So, additional studies on the accuracy of self-reported data based on large community-based probabilistic nationally representative samples are required to cover existing research gaps at the national and European levels.

Considering objective measures as a gold-standard, researchers however applied different measurement protocols and a range of cut-off values for the operational definitions of hypertension and hypercholesterolemia. Such variations may reflect differences in national reference values for some populations or a change of guidelines for clinical diagnosis over time. In any case, variability in the definitions affects the comparability of research results between studies.

Of health conditions assessed, hypertension had the most consistent operational definitions. Usually, it is measured on a single occasion several times and an average of several measurements is applied for the indicator definition. Most recent studies used JNC-5(114) definition for clinical diagnosis of hypertension, i.e. a cut-off of 140 mmHg for systolic blood pressure and 90 mmHg for diastolic blood pressure. Other studies employed the definition proposed by WHO Guidelines for the management of mild hypertension (115) that establish a 160/95 mmHg cut-off. In addition, a majority of researchers adjusted the operational definition for medication intake (33).

Definitions of objectively measured hypercholesterolemia were less consensual; studies used both fasting and non-fasting samples. It is noteworthy that most studies have used the concentration of total cholesterol as a gold-standard biomarker (10,14,96,99,116). However, Huerta and colleagues (15) and Van Eenwyk and colleagues (117) proposed more complex definitions that include additional characteristics of the lipid profile: LDL cholesterol, VLDL cholesterol, and HDL cholesterol. Although the inclusion of several specific cholesterol fractions is more in line with current clinical guidelines on cardiovascular risk assessment, implementation of this methodology on a large scale may not be cost-effective and brings additional challenges to laboratory procedures harmonization.

Furthermore, these studies applied a variety of cut-off points. Given the lack of consensus regarding the diagnosis of hypercholesterolemia studies assessing the accuracy of self-report often include sensitivity analysis, by comparing definitions of examination-based prevalence based on more than one cut-off point. In Portugal, the General Directorate of Health recommends 190 mg/dL as a therapeutic target for total cholesterol (118). This value was thus adopted in the EHES Pilot study (10). However, most of the studies adopted higher cut-off values (200 mg/dL, 220 mg/dL, 240 mg/dL, 300 mg/dL) (14,96,99,106,110,116). Given this result, further research will benefit from additional efforts to standardize health indicators definitions for lipid profile assessment at the population level.

Measurements protocols for self-reported data collection varied considerably. In addition to the difference in the mode of questionnaire administration (by interviews or self-administered), survey questions used for the definition of self-reported indicators targeted by validation studies also presented high levels of heterogeneity. An overview of survey questions used in different studies has been provided in the recent systematic review by Paalanen et al (33). The most relevant differences highlighted by the author are related to direct indication of medical diagnosis in a question formulation: "Have you been told by doctor…", the reference period for evaluation "Have you ever been told.." or "…in past 12 months". In a question formulation regarding cholesterol, most researchers used the term "high cholesterol ", but the term "hypercholesterolemia" was used less frequently.

Most frequently, studies on the accuracy of self-reported data focus on the evaluation of measurement error bias in prevalence estimates (33,103,119). Based on statistical approaches to measurement error assessment, several groups of studies can be distinguished. In the first group, self-reported and objectively measured prevalence rates are estimated for the same set of participants and measurement error bias in self-reports are computed as the absolute or relative difference between prevalence rates. The increased research comparing self-reported and objectively hypertension and hypercholesterolemia prevalence using outlined methodology allowed systematic reviews to be conducted. An excellent overview of the topic has been provided by Paalanen et al (33). According to the results of the mentioned review, bias in estimates varied by country, for hypertension the difference between the prevalence rates estimated with self-reported and objectively measured data ranged between -15.9 and +3.9 percentage points (33). In all reviewed studies except one self-reported data underestimated hypertension prevalence (33). However, the study developed by EHES Pilot Project (10) in 12 European countries has described different patterns of hypertension in females. In 7 of the 12 participating countries, the self-reported prevalence of hypertension was higher than the one objectively measured, with an absolute difference in prevalence rates between -0.9 and -15.1 percentage points. Additional research that takes a closer look at prevalence estimates for population subgroups will elucidate the differences observed in error reporting. For hypercholesterolemia bias in self-reported prevalence were even more pronounced, the difference between the prevalence rates estimated with self-reported and objectively measured data ranged between -49.6 and +11.7 percentage points (33).

The second group of studies also focused on prevalence rates, but in addition to the differences in prevalence between two data collection methods describe measurement error in terms of classification probabilities, namely sensitivity and specificity (10,14,15,66,67,96,105–108,120). Some studies also assessed overall agreement between two measurement methods using Cohen's kappa statistic.

As outlined in section 2.1.2 sensitivity of self-report correspond to the probability of correct classification among individuals who are affected by health condition. It represents a proportion of respondents who reported being diagnosed with hypertension/hypercholesterolemia among all considered as having hypertension/hypercholesterolemia according to objectively measured data. Specificity can be defined as the proportion of respondents who correctly classified themselves as not having hypertension/hypercholesterolemia among all who were healthy according to objective measurements (66).

A detailed overview of research on the sensitivity of self-reported hypertension in adult populations has been provided by Gorber et al (119) who summarized the findings of 144 observational and experimental studies developed all over the world over a 33-year period (1973 to 2006). The results showed considerable variation in sensitivity estimates and higher levels of correct reporting among women (Sensitivity range 0 - 95% compared to 0- 88% for men) (119).

The most recent systematic review by Gonçalves et al (103) covered both the sensitivity and specificity of self-reported hypertension. The authors analyzed 19 studies published between 1997 and 2015 that considered as a gold-standard objectively measured blood pressure with 140/90 cut-off. The authors reported significant variation in the sensitivity and specificity across countries. Namely, sensitivity ranged from 13% in Ghana to 92% in the USA while variations in the specificity were smaller in magnitude, from 72% in Russia to 98.9% in Australia(103). In general, lower estimates of the sensitivity were obtained for low- and middle-income countries. The results on the specificity were more consistent across the studies.

For hypercholesterolemia classification probabilities were determined by a smaller number of studies, those results are summarized in Table 2. Most of the research was conducted in the USA, the country with the longest history of HES implementation. Estimates of sensitivity varied between 27.8% in Australia and 84.9% in the USA, being lower than observed in research on hypertension. Specificity ranged from 67.3% in the USA to 99.2% in Australia. As outlined previously, comparisons between countries should be made with caution due to the remarkable heterogeneity of surveys operational definitions.

TABLE 2 SUMMARY OF ESTIMATES OF SENSITIVITY AND SPECIFICITY OF SELF-REPORTED HIGH CHOLESTEROL

	Sensitivity	Specificity
Australia		
Taylor, 2010	27.8	99.2
USA		
Bowlin, 1996	51.0	83.0
Natarajan, 2002	51.3	88.5
Bucholtz, 2018	56.9	na
Ahluwalia, 2009, 200mg/dl	71.2	78.7
Ahluwalia, 2009	84.9	67.3
Brasil		
Fantanelli, 2018	50.6	90.2
Spain		
Huerta, 2009	34.5	96.9
Korea		
Chun, 2016	46.7	95.4
Multi-country Europe		
Tolonen, 2014	30.0-32.0	93.0-95.0

Another group of studies developed within measurement error framework focused on misclassification or proportions of incorrect reporting Such studies also perform a distinction among those who suffer and who do not suffer from hypertension and hypercholesterolemia, estimating frequencies of under- and overreporting. This perspective on measurement error may be particularly useful to illustrate how many cases of hypertension and hypercholesterolemia will be missed when relying on self-reported data. Underreport can be defined as the proportion of participants failed to report hypercholesterolemia/hypertension among those with objectively measured high total cholesterol or blood pressure and can be simply estimated as 1-sensitivity.

For hypercholesterolemia, rates of underreport ranged from 40.9% reported by Man et al (121) for 40-80 yo population in Singapore to 89.0% reported by Peterson et al(68) for the Australian population aged 18 years or over. For European countries, Tolonen et al (10) reported that self-reported data failed to identify about 68-70% of individuals with total cholesterol above 190. Quite a high rate of lack of awareness of abnormal lipids levels (85%) also was estimated by ORISCAV-LUX study for the population of Luxemburg.

For hypertension, the proportion of underreporting ranged from 28% (10) to 43% and 68.4% reported by Man et al (121) and Peterson et al (68), respectively.

Several studies in addition to evaluation of the accuracy of self-reported data also investigated determinants of correct report or reporting errors. Usually, researchers addressed measurement error separately for the population groups with and without the disease. All except for one study distinguished between two possible types of an incorrect report of chronic health conditions (106). Usually, researchers addressed measurement error separately for the population groups with and without the disease, fitting two separate models: one for self-reports sensitivity/underreporting and another for self-reports specificity/overreporting.

From the methodological perspective, the research on measurement error determinants most frequently relied on a logistic regression models fitted separately for two types of reporting error using odds ratio as effect measure. Three studies applied different statistical methodologies. Ning and colleagues (66) used a multilevel logistic model with penalized quasi-likelihood estimation in order to explore contextual effects at community and regional levels. Tenkorang and colleagues (16) used a multinomial logistic regression to compare individuals' characteristics from those who "under-report" or the ones that "over-report" versus the same reference group of "correct report". Van Eenwyk and colleagues (117) used Poisson regression.

Recently it was argued in public health literature under cross-sectional design when the outcome event is not rare log-binomial and Poisson regression models are better alternatives to the logistic regression (122–124). So, the selection of appropriate link function for generalized linear models in the studies on the accuracy of self-reported data should be made taking into account the properties of available data.

Multilevel approach newly used clearly presents some advantages from a public health perspective. Literature suggests that chronic health conditions such as diabetes, hypertension and obesity are associated with socioeconomic status not only at the individual level but also at the local and the regional levels (125–128). Both contextual and individual characteristics are important to better understand social determinants of disease and their geographical distribution, particularly to design effective intervention programs targeting the most needed populations.

Research on factors associated with measurement error in self-reports of cardiovascular disease risk factors has focused mainly on the participants' characteristics. However, some studies also integrated contextual variables, such as region of residence (16,66,110) or socioeconomic condition at birth (110).

Table 3 presents a summary of factors examined in the previous research for 7 domains: demographic, socio-economic, behavioral, contextual, use of healthcare, health status and family background.

TABLE 3 VARIABLES EXPLORED IN RESEARCH ON ACCURACY OF SELF-REPORTED DATA ON CHRONIC HEALTH CONDITIONS

Domains	Variables
Demographic	Sex, Age group, Race, Ethnicity, Marital status,
Socioeconomic	Income, Education, Occupation,
Behavioral	Smoking, Alcohol consumption, Physical activity
Contextual	Region, Degree of urbanization
Health	Self-rated health, Self-reported comorbidities: diabetes, hypertension, stroke, any cardiovascular disease, depressive symptoms. Measured: Triglycerides, Glucose, Cholesterol, Obesity(BMI)
Health care	Hospitalizations, General practitioner consultations, Any medical consultations, including specialists and nurses, Last screening for diabetes, hypertension or hypercholesterolemia
Family	Household size, Socioeconomic status at birth (born in poor family), Family history of cardiovascular disease

Prior research documented the association between self-report accuracy of cardiovascular disease risk factors and participants' characteristics. Females were more likely to report their hypertension status correctly in several studies (15,66,67). While for hypercholesterolemia in the majority of the research settings no statistically significant associations with sex were observed and lower sensitivity was observed for females in Korea (96).

Most of the studies reported an increase of self-report sensitivity with age for hypercholesterolemia (15,99) and for hypertension(66,67,96).

Regarding health behaviors, the association between smoking status and accuracy of hypercholesterolemia self-reports was observed in a single study (96).

Several studies sought to determine whether the accuracy of survey self-reports is related to socioeconomic status. There is some evidence regarding an educational gradient on incorrect self-reports of hypertension in China (66) and Taiwan (120), where individuals with higher levels of education are more likely to report their health condition more accurately. Findings of the Korean National Health Examination and Nutrition Survey (96) suggest that the likelihood of incorrect reports of hypercholesterolemia is higher for individuals with higher education. Similar findings were reported by Choi and Cawley (69). Moreover, higher sensitivity of self-reported cardiovascular disease risk factors was observed for individuals with higher income (66,96).

There is limited evidence regarding the role of socioeconomic factors in self-reported data accuracy for the European population. Johnston (129) has reported an inverse educational gradient for undiagnosed hypertension in England; less educated individuals were more likely to have a disease, but at the same time were not aware of their health condition, and as such, more likely not to report it. Tompkins (107) found no association between education levels, occupation and self-reports of hypercholesterolemia.

Although not consistent, these results suggest that measurement error in self-report may be differential among individuals from different socioeconomic groups. In such case, survey estimates based on self-reports may be more distant from “true” values for some population subgroups than for others. Differential measurement errors can jeopardize group comparisons. Limitations of self-reported data were also recognized in research on health inequalities (130), however, few studies performed a formal assessment of measurement error on inequalities estimates. Mackenbach et al (131) showed that reporting error in diabetes and heart disease varied by level of education and self-reported data underestimate inequalities in the prevalence of disease in the Netherlands. Vellakkal et al (132) demonstrated that socioeconomic inequalities in hypertension in low- and middle-income countries may be underestimated, or even have an opposite direction when self-reported data is used. In the USA, according to NHANES data, self-reporting overestimated the educational disparities in hypercholesterolemia and underestimated disparities in hypertension when comparing prevalence rates by education (69). In general, studies on this topic are scarce, to our best knowledge, no studies has been published yet at the national and EU level.

Survey statisticians and epidemiologists are seeking efficient methods for correction of measurement error in prevalence estimates and outcome-exposure associations during the data analysis stage. Despite serval methodological options currently available, in practice, still few researchers test the robustness of study findings in presence of measurement error.

Regarding hypertension and hypercholesterolemia, in a literature search I identified only four research papers that attempt to perform the correction. Three of them applied so-called regression calibration and one illustrated the application of multiple imputation for measurement error correction for self-reported hypertension. For both correction methods, the researchers used individual-level data from external samples to perform adjustments. Misclassification model was established for one survey that combines self-reported and objectively measured data and when transferred to another survey that required correction. A regression-based approach with sensitivity and specificity correction for categorical outcomes was proposed by Mentz et al (34) and later applied by Yi et al (73) and Ni et al. Schenker et al (12) applied a multiple imputation approach to correct estimates from the USA National Health Interview Survey using data on hypertension objectively measured by NHANES. Both methods successfully corrected measurement error bias in prevalence estimates, including estimates in population subgroups. Schenker et al (12) also reported successful correction of bias in regression coefficients in a logistic regression model with the misclassified outcome. Additional research is required to investigate the feasibility of measurement error correction.

Despite observed heterogeneity of research methods, it is clear that there is variation between self-reported and objectively measured hypercholesterolemia and hypertension in health surveys. The research on survey self-report measurement error quantification and correction was developed mostly in countries with a long tradition of HES implementation. Given the expressed need for an integrated and sustainable system of health indicators comparable over time and countries for continuous health monitoring in Europe (4) further research on the comparability and accuracy of self-reported data is required. Hence, from a public health perspective, it is important to understand for Portuguese population:

1. whether measurement error is epidemiologically relevant, i.e. to what extend self-reported data under- or over-estimates the prevalence of risk factors for cardiovascular disease;
2. whether measurement error is differential between population subgroups;
3. what are the implications of measurement error on the comparability of self-reported disease status across demographic and socioeconomic groups;
4. how to control the measurement error during survey planning and implementation; and
5. whether it is feasible to account for measurement error using statistical approaches.

3 Research objectives

In this Chapter, I outline the main and specific thesis objectives.

The main aim of this thesis is to contribute to the improvement of health information available for monitoring of two cardiovascular disease risk factors (hypertension and hypercholesterolemia).

This will be achieved by evaluating to what extent objectively measured and self-reported data differ for the Portuguese population and, if so, describe quality control measures at the data collection stage and propose a methodology to adjust for measurement error at the data analysis stage.

Specific objectives include:

Objective 1. To estimate self-reported and examination-based prevalence of hypertension and hypercholesterolemia and examine to which extent the prevalence rates differ between the two data collection methods (i.e., determine direction and magnitude of bias in self-reported prevalence estimates due to measurement error); (hereby referred to as “Quantification”);

Objective 2. To evaluate association of measurement error in self-reported hypertension and hypercholesterolemia with demographic and socioeconomic characteristics of survey participants (hereby referred to as “Determinants”);

Objective 3. To estimate socioeconomic inequalities in the distribution of hypertension and hypercholesterolemia in the Portuguese population using self-reported and objectively measured data and assess the effect of measurement error in self-reported information on estimates of socioeconomic inequality indicators (hereby referred to as “Quantification”).

Objective 4. To describe quality control measures during survey planning and implementation aimed to reduce measurement error (hereby referred to as “Control”);

Objective 5. To apply statistical methods to correct measurement error in self-reported data to accurately estimate the prevalence of hypertension and hypercholesterolemia, and measures of exposure-outcome associations (hereby referred to as “Correction”).

The studies presented in this thesis attempt to cover three main research gaps, both at the national and European level. First, developed research allows the quantification of existing differences between self-reported and objectively measured prevalence of hypertension and hypercholesterolemia using a large probabilistic sample representative of the Portuguese population, thus addressing an important research gap in the country-level.

Secondly, study of the effect of measurement error in self-report on estimates of socioeconomic inequality indicators will be innovative both at the national and European level.

Finally, statistical approaches to correct residual measurement error in self-reported data are applied. I investigate a feasibility of measurement error correction in prevalence estimates and in exposure-outcome associations at the data analysis phase, using internal and external validation data sources.

4 Methods

In this section I present methodological approaches used to objectives outlined in Chapter 3, including study settings, data sources used in research and detailed description of statistical analysis techniques for each particular study.

4.1 Study design

To achieve the established objectives four observational, cross-sectional epidemiological studies with descriptive and analytical components were developed. Correspondence between objectives outlined in previous chapter and research work is presented in Table 4.

TABLE 4 RESEARCH WORK CORRESPONDING TO EACH OBJECTIVE

Objective	Study	Research paper
Objective 1 Quantification	Study 1	Kislaya el al (2019) Differential self-report error by socioeconomic status in hypertension and hypercholesterolemia: INSEF 2015 study. European Journal of Public Health, Vol. 29, No. 2, 273–278, https://doi.org/10.1093/eurpub/cky228
Objective 2. Determinants		
Objective 3. Quantification	Study 2	Kislaya et al (2019) Do self-reported data accurately measure health inequalities in risk factors for cardiovascular disease? International Journal of Public Health, doi:10.1007/s00038-019-01232-1.
Objective 4. Control	Study 3	Kislaya et al Collecting valid and reliable data: fieldwork monitoring strategies in a Health Examination Survey. Portuguese Journal of Public Health, DOI: 10.1159/000511576

Objective 5. Correction	Study 4	Kislaya et al Combining self-reported and objectively measured survey data to improve hypertension prevalence estimates: Portuguese experience, Arch Public Health. 2021 Apr 8;79(1):45. doi: 10.1186/s13690-021-00562-y.

4.2 Target population

In all the studies developed the target population included individuals aged between 25 and 74 years old, residing in Portugal (mainland or autonomous regions) in private households in 2014-2015, and who speak/able to follow an interview in Portuguese.

The exclusion criteria were: 1) being a resident in collective households or institutions homes for the elderly, nursing homes, psychiatric hospitals, military housing units, prisons, or others; 2) unable to understand an interview in Portuguese; 3) being unable to give informed consent or to complete the survey due to physical disability, mental illness, dementia or other condition that seriously affects the ability of understanding and response accuracy.

4.3 Data sources

I used data from two Portuguese population-based surveys: the first Portuguese National Health Examination Survey (INSEF) and the National Health Interview Survey (INS2014), the Portuguese component of the European Health Interview Survey (EHIS), second wave. For study 1 to 4 INSEF data was used. For Study 5 we used both INSEF and INS2014 data.

4.3.1. Portuguese National Health Examination Survey (INSEF)

The INSEF was developed by the National Health Institute Dr. Ricardo Jorge (INSA) in partnership with the Norwegian Institute of Public Health (NIPH), and in collaboration with the Regional Health Administrations (ARS) and the Regional Secretariats for Health of the Autonomous Regions of the Azores and Madeira in 2013-

2016 as part of the project "Improvement of epidemiological health information to support public health decision and management in Portugal. Towards reduced inequalities, improved health, and bilateral cooperation". INSEF combined self-reported data (health interview) with collection of objective health measures through physical examination and blood collection.

Sampling procedure

To obtain the INSEF sample, a two-stage cluster sampling stratified by region and type of urban area (TUA) (rural versus urban and semi-urban) was performed. In the first sampling stage, in each stratum (region/TUA) Primary Sampling Units (PSU) were selected. The PSUs were geographical areas corresponding to the catchment areas of Health Centres of the National Health Service.

Health Centres catchment area were defined as PSU for various reasons: i) logistics, ii) issues associated with INSEF survey features, iii) observational costs, iv) catchment areas include a well-defined served population, v) catchment areas population can be adequately characterized by sex and age group, using official statistics, vi) conduction of the survey in primary care settings increases survey credibility and survey participation.

National territory was divided into 386 PSU. Geographical information systems were employed to define PSUs' boundaries and evaluate PSUs' spatial continuity. Minimum size required for a PSU was established at 1000 individuals (residents). The selection of PSUs was performed with a probability proportional to the size of the 25-74 years old population resident in the catchment area, i.e PSU with the higher number of residents had the higher probability of being selected.

In the second stage, in each selected PSU, individuals (Secondary sampling units - SSU) were selected by simple random sampling from the National Register of Health Care Users that covers approximately all resident population. The selection of SSU within PSU was performed as close as possible to the time of fieldwork in the respective PSU (usually one month before the start of fieldwork) to ensure up to date information in a sampling frame.

Sample size and precision

INSEF sample size was established at 600 individuals in each of the seven regions, in order to estimate an expected prevalence of 50%, with a precision of 5% for

95% confidence level, considering a design effect of 1.5 at regional level. Homogeneous allocation of the sample by region was used to obtain comparable estimates of health indicators with similar precision for all seven regions. At national level, the minimum sample size was set as 600 per region which accounts for a total of 4200. Results of the pilot survey developed in Portugal (ref) and other HES in European countries (133) indicated an expected participation rate of approximately 40%. Therefore, oversampling factors according to this value were applied to inflate the number of individuals to invite in each PSU .

Overall, final INSEF sample consisted of 12289 individuals of which 4911 effectively participated in survey (participation rate of 43.9%) (134). Minimum required sample size was reached in all seven regions. At national level, achieved sample size allows to estimate an expected prevalence of 50%, with an absolute margin of error of ±1.7% at a confidence interval of 95%, considering a design effect of 1.5. Precision of estimates for other possible values of expected prevalence are illustrated in Table 5.

TABLE 5 EXPECTED ACCURACY FOR DIFFERENT EXPECTED PREVALENCE LEVELS, ASSUMING A SAMPLE SIZE OF 4911, A STUDY DESIGN EFFECT OF 1.5 AND A CONFIDENCE INTERVAL OF 95% FOR THE ESTIMATED PREVALENCE

Expected prevalence	5%	10%	20%	30%	40%	50%
Absolute margin of error	+/-0.7%	+/-1%	+/-1.4%	+/-1.6%	+/-1.7%	+/-1.7%

Data collection

INSEF data collection was performed in February-December 2015 by 24 regional teams, each consisting of 1 administrative officer (performed recruitment of participants), 2 nurses (performed the interviews and physical examinations) and 1 laboratory technician or nurse (performed the blood collection and processing).

To ensure the standardization of survey procedures, all regional team members (n=117) participated in the training program. Training sessions were developed in each region, as close as possible to the time of data collection, with a duration of 21 to 28 hours. Additionally, fieldwork staff received manuals with step-by-step description of the survey procedures to which they were assigned.

Self-reported data was collected by computer assisted personal interview, using research electronic data capture software REDcap (135); in-paper forms were used to register objective health measures and blood collection and processing.

Measurements

All measurements in INSEF were conducted using a standardized measurement protocol based on the recommendations of the European projects: Feasibility of a European Health Examination Survey and the European Health Examination Survey pilot Joint Action (20,136–138).

Blood pressure was measured using an automated sphygmomanometer OMROM M6. According to the study protocol, measurements were taken on the participant right arm. In case of amputation, cast, open wounds/sores, rash or malformation preventing to place the cuff properly on the right arm, measurements were taken on the left arm. Cuff of suitable size was selected for each participant based on his/her arm circumference. Measurements were taken in the sitting position following a five-minute rest period. Three consecutive blood pressure measurements were taken with a one-minute interval between each measurement. All the results were read in mmHg. The mean value of the 2nd and the 3rd measurements of systolic and diastolic blood pressure were used as blood pressure value.

According to the INSEF protocol, venous blood samples were collected from all participants except those with anaemia or other chronic illnesses, which restricts taking blood samples. Up to two attempts of blood withdrawal were made. Total cholesterol was measured on the serum isolated from the non-fasting blood samples using the cholesterol esterase/oxidase method. Total cholesterol concentration was expressed in conventional units of measurement (mg/dL) as these are the units commonly used in Portugal.

Survey Quality control

Supervision of INSEF fieldwork aimed to assure the proper application of recruitment, interviewing and measuring protocols and hence reduce total survey error. A set of indicators was monitored continuously during the fieldwork. According to recommendations for HES in Europe (138), retrospective data quality assessment was performed using quality scores (139–141) for each survey procedure at the end of the fieldwork.

All 12 regional laboratories involved in sample processing participated in the National Programme for Laboratory Quality Assurance (PNAEQ). At six different time points during fieldwork, sets of two control samples of known total cholesterol concentration were analyzed by participating laboratories. The results of external quality control were satisfactory, since all participating laboratories presented values of Standard Deviation Index (SDI) within ±3.

Time of interview, proportion of item non-response and consistency of response were monitored for questionnaire data. In case of inconsistencies identified in survey response participants were re-contacted via telephone for clarification and correction.

4.3.2. National Health Interview Survey (INS 2014)

The National Health Interview Survey (INS2014) was developed by Statistics Portugal in collaboration with INSA as a part of the second wave of the European Health Interview Survey under the Regulation (EC) No 1338/2008 of the European Parliament and of the Council on Community statistics on public health and health and safety at work and Commission Regulation (EU) No 141/2013.

The INS 2014 target population included individuals aged 15 years old and over residing in Portugal, excluding those residing in collective households or institutions. The survey was designed to be representative at national and regional level.

Sampling procedure

The INS2014 sample was selected using systematic three-stage cluster sampling stratified by region NUTs II (1989 and 2002). At the first stage, geographical areas corresponding to the Census Sections were selected with probability proportional to size. At the second stage, within each selected census section, households were selected by systematic sampling using National Dwelling Register updated according to the Census 2011. At the final stage, one individual per household was selected using the “last birthday” rule (142). More detailed description of survey sampling has been published elsewhere (142).

Sample size and precision

INS 2014 sample size was established at 22538 individuals in order to estimate key indicators of health status and health determinants (self-perceived general health, prevalence of diseases and chronic conditions, smoking and alcohol consumption) with a relative standard error (coefficient of variation) not exceeding 12% for Autonomous Region of Madeira and 10% for the remaining NUTs II 1989 and NUTs II 2002 regions.

The effectively sample size reached was of 18204, corresponding to a response rate of 80.8% (142).

In this thesis we used a INS2014 subsample, applying a restriction to 25-74 years old (n= 13937) to corresponded to INSEF study population. This sample size allows estimating an expected prevalence of 50%, with an absolute margin of error of ±1.0% at a confidence interval of 95%, considering a design effect of 1.5 at national level. Alternative values of expected prevalence and corresponding precision estimates for the sample size achieved are presented in Table 6.

TABLE 6 EXPECTED ACCURACY FOR DIFFERENT EXPECTED PREVALENCE LEVELS, ASSUMING A SAMPLE SIZE OF 13937, A STUDY DESIGN EFFECT OF 1.5 AND A CONFIDENCE INTERVAL OF 95% FOR THE ESTIMATED PREVALENCE

Expected prevalence	5%	10%	20%	30%	40%	50%
Absolute margin of error	+/-0.4%	+/-0.6%	+/-0.8%	+/-0.9%	+/-1%	+/-1%

Data collection

INS2014 fieldwork took place in September-December 2014. Survey data were collected combining two data collection modes: face-to face interviews (93.8%) and self-administered web questionnaires (6.2%). Interviews were performed during the households visits by 167 interviewers, experienced with health and social surveys. All the staff involved in data collection process completed a training program of 14 h on survey procedures and received a training manual with detailed interview guide.

Survey quality control

Supervision of INS2014 fieldwork was performed by the Statistics Portugal staff following Eurostat's guidelines for the survey implementation and quality control. This aimed to minimize total survey error. Proportion of unit and item non-response and interview duration were continuously monitored. To verify whether interviews were actually undertaken a random sample of survey participants (3%) was re-contacted via telephone. Data checks were performed to assess the plausibility and consistency of survey response.

4.4 Operationalization of variables

Data on objective measurements of blood pressure and total cholesterol as well as self-reported socio-demographic information, information on health status, medication uptake and healthcare use was used. The full list and description of collected variables available for this study is provided in Appendix B.

4.4.1. Outcome variables

Examination-based outcome variables

The definitions proposed by the European Health Examination Survey guidelines (143) were employed for examination-based prevalence. Examination-based prevalence of hypertension and hypercholesterolemia was based on objective measures of these health conditions corrected for the use of medication. The rational for this choice is that medicated individuals, who have their health condition under control due to medical treatment, should be considered as having the condition regardless of their current blood pressure/cholesterol values. These individuals are not ”disease free”, as they still have the condition that is under control. As such, these individuals should be included in the numerator of the prevalence rate. The inclusion of self-reported medication in examination-based definition is a common practice in HES performed worldwide(10,14,66,67,99)

Information of medication intake was obtained from two questions: “during the past two weeks, have you used any medicines that were prescribed for you by a doctor?” and if yes, ”were the medicines for hypertension / hypercholesterolemia? (Yes/No)”.

Examination-based hypertension was defined as:

(i) systolic blood pressure of at least 140 mmHg, or

(ii) diastolic blood pressure of at least 90 mmHg, or

(iii) intake of antihypertensive medication prescribed by a doctor, in the two weeks prior to the interview.

Hence, the examination-based hypertension prevalence will correspond to the proportion of individuals with hypertension objectively determined from all the survey participants with available data on both the measured and self-reported condition. The cuf-off 140/90 mmHg was selected based on the EHES recommendations, JNC-5 (114) definition for clinical diagnosis of hypertension and national guidelines.

Examination-based prevalence of hypercholesterolemia was defined as the proportion of individuals with a total serum cholesterol concentration of at least 190 mg/dL or who reported using prescribed lipid-lowering medication in the two weeks prior to the interview among all the survey participants with available data on both the measured and self-reported condition. The cut-point of 190mg/dl (= 5.0 mmol/l) was selected based on the EHES recommendations and also on Portuguese and EU level treatment guidelines (118,143–145) in place at the time of data collection.

Self-reported outcome variables

For self-reported prevalence, the definitions are in line with the guidelines proposed for the second wave of the European Health Interview Survey (146). Self-reported prevalence was given by the proportion of individuals who reported the respective health condition among all survey participants with available information on that specific health condition.

In INSEF survey self-reported prevalence of hypertension and hypercholesterolemia was defined based on the following questions:

- “Do you have any of the following longstanding diseases or conditions (which have lasted, or are expected to last, for 6 months or more): High blood pressure or hypertension; hypercholesterolemia? (Yes/No)”; and if yes,
- “Were these conditions diagnosed by a medical doctor? (Yes/No)”. Individuals were considered to have self-reported hypertension/hypercholesterolemia if they answered positively to both questions.

In INS2014 a slightly different question was used to define self-reported prevalence of hypertension:

- “During the past 12 months, have you had any of the following disease or conditions? High blood pressure/Hypertension (Yes/No). Consider disease/conditions even if the symptoms were not present due to a medical treatment”. Individuals who answered positively were considered hypertensive.

It is of note that the definition of self-reported prevalence proposed by guidelines for the second wave of the European Health Interview Survey does not account for medication intake. It included questions on chronic diseases/conditions but no questions on medication intake for specific conditions. Participants were asked in general if they were taking any prescribed medication, but no further details on the medication taken were asked. Hence, with INS2014 available data on medication, it is not possible to link the medication to specific disease/condition, as it can be done in the examination-based survey INSEF.

4.4.2. Exposure variables

Socioeconomic status of participants was measured by education and equivalised disposable household income.

Education refers to the highest level of education successfully completed. It was measured by the question:

“What is the highest level of education that you have completed successfully?” and categorized into four classes using the International Standard Classification of Education (ISCED 2011)(147).

The equivalised disposable household income refers to a total income of a household available for spending or saving after tax and other deductions, adjusted for household composition . The modified OECD equivalence scale was used; to account for the household composition the following weights were attributed for each household member:

- 1 for the first household member aged 14 years and older;
- 0.5 for the second and each subsequent household member aged 14 years and older;
- 0.3 for each household member aged less than 14 years old.

The equivalised disposable household income was calculated dividing the total household income by the sum of weights for all the household members. After adjustment for household composition, income was categorized in quintiles, with approximately 20% of individuals in each, from the lowest income coded as 1st quintile (1Q) to the highest coded as 5th quintile (5Q).

As outlined previously, description of other exposure variables available for this study is provided in Appendix B.

4.5 Statistical analysis

4.5.1 Weighting

Both INSEF and INS2014 were developed under a complex sample design, and we intended to obtain nationally representative results. Therefore, sampling weights were used in data analysis, as recommended in literature (148,149), for all studies, except Study 3.

Sampling weights express the number of individuals in the target population represented by each individual in the sample. Sampling weight are used to account for several survey features, such as the differential probabilities of PSU and SSU selection, unit nonresponse and differences between the final sample distribution and the target population distribution. Further details on the sampling weight calculation for INSEF and the INS2014 samples are provided elsewhere (142,150).

4.5.2. Study 1.

Study 1 was based on the INSEF data and aimed to examine whether the measurement error in self-reported data is epidemiologically relevant and identify associated factors.

The prevalence of hypertension and hypercholesterolemia at national level and stratified by participants' characteristics was estimated using examination-based and self-reported data.

Absolute and relative differences between self-reported and examination-based prevalence were calculated for the overall sample and for each population subgroup under study.

Sensitivity and specificity for self-reported data were computed using the examination-based results corrected for current medication intake as the gold-standard measure.

Sensitivity of self-report was defined as the percentage of respondents who reported being diagnosed with hypertension/hypercholesterolemia among all considered as having hypertension/hypercholesterolemia according to examination-based definitions (66).

Specificity of self-report was defined as the percentage of respondents who reported not having hypertension/hypercholesterolemia/diabetes among all who were classified as not having the health condition according to examination-based definitions (66).

Sensitivity and specificity of self-reported hypertension and hypercholesterolemia were compared between population subgroups using design adjusted Rao-Scott version of chi-square test (151,152).

Overreport of hypertension or hypercholesterolemia, defined as the proportion of participants who had reported hypertension/hypercholesterolemia in the interview but were classified as not having hypertension/hypercholesterolemia according examination-based definitions, was calculated as 1-Specificity.

Underreport of hypertension or hypercholesterolemia, defined as percentage of participants who fits the examination-based definitions but fail to report their health condition, was computed as 1-Sensetivity.

To identify factors associated with underreport of hypertension and hypercholesterolemia Poisson regression models were used. As a measure of association adjusted prevalence ratios (PRs) were estimated. The set of independent variables to be included in the models was based on the results of the literature review performed as part of this work, their availability in the INSEF survey database and statistical criteria

4.5.3. Study 2.

Study 2 was based on the INSEF data and aimed to examine the effect of self-report measurement error on outcome-exposure associations. Self-reported and examination-based hypertension and hypercholesterolemia were considered as outcome variables and education level as the main exposure. As measures of association I considered regression-based health inequality indicators, the Relative Inequality Index (RII) and the Slope Inequality Index (SII).

Absolute and relative educational inequalities were estimated for four population subgroups: 25-49 female, 25-49 male, 50-74 female, 50-74 male using self-reported and examination-based data.

Relative inequalities in the distribution of hypertension and hypercholesterolemia were measured by the RII and absolute inequalities were measured by the SII. To obtain RII and SII inequality measures, I first converted the exposure variable, education level, to the so-called “ridit” score for each sex-age stratum. To compute ridit, individuals within each stratum were ranked according to their socioeconomic position (from the highest education level to the lowest) and for each educational group the midpoint of cumulative distribution was attributed, as previously described by Ernstsen et al (153)and Mackenbach & Kunst (154)

As two types of measurement of hypertension and hypercholesterolemia were available for each study participant, I used Generalized Estimating Equations (GEE) that allows to estimate and formally compare absolute and relative inequalities between self-reported and examination-based data for each outcome.

For stratified analysis in each population subgroups, GEE models included age group, ridit, type of measurement (self-reported or examination-based) and the ridit*type of measurement interaction.

$$g(Y) = \beta_0 + \beta_1 ridit + \beta_2 type\_of\_measurement + \beta_3 ridit * type\_of\_measurement + \beta_4\, age\_group + error$$

Overall sample models were additionally adjusted for sex.

I used Possion model with logarithmic link function ($g(Y) = \log(Y)$) to estimate RII, while to estimate SII I employed Poisson model with identity link ($g(Y) = Y$).

I consided self-reported data as "baseline" measurement, so exponential of the model coefficient for ridit (β_1) in the log-link Poisson model was considered the RII estimate for self-reported data (RIIsr):

$$RII_{sr} = e^{\beta_1}.$$

Ridit score can be interpreted as a continuous variable that takes values between zero and one, a value of zero corresponds to the highest possible educational level and a value of one to the lowest possible level in educational hierarchy. Therefore, RII represents the prevalence rate ratio between those with rank one (the lowest level of education) and those with rank zero (the highest level of education). RII values greater than one indicate an inverse educational gradient in hypertension/hypercholesterolemia, or in other words, that prevalence rates of hypertension/hypercholesterolemia are higher among those with lower levels of education.

Exponential of the model coefficient for ridit*type of measurement interaction (β_3) represents the ratio of examination-based and self-reported relative inequality measures:

$$e^{\beta_3} = \frac{RII_{eb}}{RII_{sr}}.$$

When RIIeb/ RIIsr ratio is statistically significant and greater than one, RIIeb is greater in magnitude than RIIsr, indicating that self-reported data underestimate educational inequalities in the outcome prevalence. Statistically significant RIIeb/ RIIsr values below one indicate that self-reported data overestimate educational inequalities in the outcome prevalence.

Examination-based RII (RIIeb) was calculated using the previously describe model as follows:

$$RII_{eb} = e^{(\beta_1+\beta_3)}.$$

A similar approach was used for SII. As it was estimated by an additive GEE Poisson model (with identity link function, $g(Y) = Y$), SII represents the prevalence rate difference between those in the top and bottom of ridit ranking. Positive SII indicates higher hypertension/hypercholesterolemia prevalence among less educated. Coefficient of ridit*type of measurement interaction represents the difference between examination-based and self-reported SII:

$$\beta_3 = SII_{eb} - SII_{sr}.$$

Positive values of this expression mean that SIIeb is higher than SIIsr, indicating that self-reported data underestimate absolute inequalities in hypertension/hypercholesterolemia prevalence and negative values overestimate it.

4.5.4. Study 3.

Study 3 was based on INSEF paradata collected during survey fieldwork implementation to evaluate overall survey performance and detect deviations from the survey protocol. To assure the proper application of interviewing and measuring protocols and hence reduce measurement error, a set of Key Performance Indicators (Table 7) was established based on EHES guidelines, World Health Organization MONICA Project recommendations and previous practical experience of data gathering for EHES pilot in 2009–2010 (155).

TABLE 7 . KEY PERFORMANCE INDICATORS FOR FIELDWORK MONITORING, INSEF 2015

Survey component	Key Performance Indicators	Units	Target Performance
Recruitment	Contact, cooperation and participation rate	Pilot and PSUs for all the regions (PSU 1 through PSU7)	40% participation rate

Physical exam	Average time for blood pressure measurement	Pilot and PSUs for all the regions (PSU 1 through PSU7)	Minimum duration of 8 minutes
	Verification of the last digit distribution for height	Pilot and PSUs for all the regions (PSU 1 through PSU7)	Uniform
Blood collection and processing	Proportion of haemolysed serum samples	PSUs for all the regions (PSU 1 through PSU7)	Below 5%
Health questionnaire	Average interview time	PSUs for all the regions (PSU 1 through PSU7)	Minimum of 15 minutes
	Average rate of items missing for selected questionnaire areas	Overall	Below 5%

Contact, cooperation and participation rates were computed according to HES recommendations (137) as follows:

$$\textbf{Participation rate} = \frac{N^{\text{o}}\ participants}{N^{\text{o}}\ eligible + N^{\text{o}}\ unresolved};$$

$$\textbf{Contact rate} = \frac{N^{\text{o}}\ eligible}{N^{\text{o}}\ eligible + N^{\text{o}}\ unresolved};$$

$$\textbf{Cooperation rate} = \frac{N^{\text{o}}\ participants}{N^{\text{o}}\ eligible},$$

where eligible individuals were selected individuals successfully contacted and belonging to the target population when the survey took place; participants were eligible individuals who performed at least one physical measurement successfully, accepted two attempts of blood sample draw and answered at least 50% of the questions during the interview; and unresolved were individuals who were not reached.

Relative frequencies and means were used to monitor other KPIs along the timeline of the fieldwork period. The results were presented by sets of seven observations throughout the survey period to evaluate trends within the survey - each group is composed of one PSU from each one of the seven health regions, corresponding to the different health regions first, second, third, fourth, fifth, sixth and seventh PSUs of each.

4.5.5. Study 4.

Study 4 aimed to investigate the feasibility of measurement error correction in self-reported hypertension data under a multiple imputation approach (MIME) using internal and external validation data on objectively measured outcome. It used both INSEF and INS2014 data. This study was limited to hypertension, as hypercholesterolemia was not assessed in the INS 2014 survey.

Given the binary nature of the outcome variables (hypertension), for measurement error correction I applied logistic regression imputation method under assumption of monotone missing data patterns.

First, to explore the relationship between the objectively measured outcome Y, self-reported error-prone version of the outcome (X) and the set of other covariates ($Z_i \in \boldsymbol{Z}$), a logistic regression model (Equation X) was fitted on the INSEF sample:

$$P(Y = 1|X, Z_i) = \frac{\exp(\beta_0 + \beta_1 X + \sum_{i=2}^{I} \beta_i Z_i)}{1 + \exp(\beta_0 + \beta_1 X + \sum_{i=2}^{I} \beta_i Z_i)}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_I)$ represents a vector of regression parameters.

Based on recommendations of the multiple imputation literature, as potential covariates (Zi) for the model I considered 3 groups of variables: i) those related to the measurement error in the Portuguese context, ii) those identified in the literature as risk factors for hypertension and iii) survey design variables to account for stratification and weighting. Potentially relevant variables were identified based on Study 1.

For each fitted model, I computed the area under receiver operating curve (ROC) to measure models´ predictive ability. To assess models performance I compared area under ROC; the model with the highest area was selected to be the final imputation model.

As a second step, the imputation model was tested on INSEF sample using 1000 simulations. In each of 1000 simulations INSEF sample was spitted into two parts. The first part was assumed to be a validation sub-sample with available data on both true outcome (examination-based) and its error-prone version (self-reported) and was used to fit the imputation model. For the remaining sub-sample (training) I assumed that “examination-based” outcome values were missing and the model selected in the first step was used to multiply impute outcome values in this sub-sample.

Augmented regression option was used to deal with potential perfect predictions situations. Set of 100 imputations was created in each simulation run. For each imputation k= 1, 2, …, 100 I draw a set of regression coefficients $\boldsymbol{\beta^k}$ from the posterior predictive distribution of regression parameters $\boldsymbol{\beta}$. For each participant in a training sub-sample using previously dawn regression parameters $\boldsymbol{\beta^k}$ I computed an expected probability of $Y^k = 1$:

$$P\left(Y^k = 1\right) = p^k = \frac{\exp(\beta_0^k + \beta_1^k X + \sum_{i=2}^{I} \beta_i^k Z_i)}{1 + \exp(\beta_0^k + \beta_1^k X + \sum_{i=2}^{I} \beta_i^k Z_i)}.$$

Values of missing “examination-based” outcome (Y^k) in each of k imputations were drawn from the Bernoulli distribution with a probability p^k.

For each of generated imputation k dataset I estimated prevalence rates and prevalence ratios. The resulting estimates were combined using Rubin´s rules to account for the variability introduced through the imputation procedure. Based on 1000 simulations results I constructed empirical distributions of prevalence rates and prevalence ratios.

In the next step, descriptive statistics (absolute and relative frequencies for categorical variables) to characterize INS2014 and INSEF samples and computed objectively measured and self-reported hypertension prevalence were obtained. I compared participants´ characteristics and prevalence of self-reported hypertension between two surveys using design adjusted Rao-Scott version of chi-square test (151,152).

Finally, the previously developed imputation model was used to impute objectively measured hypertension for INS2014 participants. Performance of MIME for prevalence ratios and prevalence rates corrections was assess in terms of bias, SE and MSE.

4.6 Sensitivity analysis

The definition of examination-based hypercholesterolemia is not consensual in research and medical practice (33). To evaluate the robustness of our findings regarding hypercholesterolemia, alternative definition for examination-based prevalence of hypercholesterolemia were used in sensitivity analysis:

The proportion of individuals with a total serum cholesterol concentration of at least 200 mg/dL or who reported using prescribed medication to control blood lipid levels in the two weeks prior to the interview among all the survey participants.

In addition, to investigate robustness of study findings, operational definition of self-reported hypertension and hypercholesterolemia included information on medication uptake in previous two weeks.

4.7 Statistical software

All data management and analysis was performed using STATA statistical software. For the Study 1 Stata 11.2® version was used, for the remaining studies Stata 15.1® version was employed (156). To account for complex sample design, [svy] module routines were used.

4.8. . Ethical considerations

This research was developed in compliance with existing the regulations on data protection and human participation in biomedical research. I used anonymous data from two population-based surveys, the INSEF and INS2014.

The INSEF received clearance from the National Commission for Data Protection, by the Ethical Committees of INSA and ethical committees of all project partners. Participation in the INSEF study was voluntary, each participant provided a written informed consent. The confidentiality of participants was assured through suppression of all personal information allowing direct or indirect identification from the database after the ending of recruitment and data collection process.

The INS2014 was developed in accordance with the principles of statistical confidentiality under the Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics. Released individual-level data was anonymized in a way that does not allow either direct or indirect identification of sampling units.

Scientific protocol of the thesis (Including Studies 1-4) was approved by INSA Ethical Committee.

5 Results

In this Chapter I present the main results of developed studies. Chapter is organized into four section, each dedicated to specific research paper.

5.1. Differential self-report error by socioeconomic status in hypertension and hypercholesterolemia: INSEF 2015 study.

The first research paper was published in 2019 in the European Journal of Public Health. It discuss magnitude and direction of bias in self-reported hypertension and hypercholesterolemia: INSEF 2015 study and factors associated with underreporting of selected health conditions by Portuguese Population

European Journal of Public Health, 1–6

doi:10.1093/eurpub/cky228

Differential self-report error by socioeconomic status in hypertension and hypercholesterolemia: INSEF 2015 study

Irina Kislaya [1,2], Hanna Tolonen[3], Ana Paula Rodrigues[1], Marta Barreto [1,2], Ana Paula Gil[4], Vânia Gaio [1,2], Sónia Namorado [1,2], Ana João Santos [1], Carlos Matias Dias [1,2], Baltazar Nunes [1,2]

1 Departamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge, IP, Lisboa, Portugal
2 Centro de Investigação em Saúde Pública, Escola Nacional de Saúde Pública, Universidade NOVA de Lisboa, Lisboa, Portugal
3 Department of Public Health Solutions, National Institute for Health and Welfare (THL), Helsinki, Finland
4 Interdisciplinary Centre of Social Sciences (CICS.NOVA) NOVA School of Social Sciences and Humanities, Universidade NOVA de Lisboa, Lisboa, Portugal

Correspondence: Irina Kislaya, Departamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge, IP, Avenida Padre Cruz, 1649-016 Lisboa, Portugal, Tel: (+351) 217 526 492, Fax: (+351) 217 526 400, e-mail: irina.kislaya@insa.min-saude.pt

Background: This study aimed to compare self-reported and examination-based prevalence of hypertension and hypercholesterolemia in Portugal in 2015 and to identify factors associated with the measurement error in self-reports. **Methods:** We used data from the Portuguese National Health Examination Survey ($n = 4911$), that combines personal interview, blood collection and, physical examination. Sensitivity and specificity of self-reported hypertension and hypercholesterolemia were calculated. Poisson regression was used to estimate prevalence ratios (PRs) of underreport of hypertension and hypercholesterolemia according to sex, age, socioeconomic status (education and income) and general practitioner (GP) consultation in the past year. **Results:** Sensitivity of self-reports was 69.8% for hypertension and 38.2% for hypercholesterolemia. Underreport of hypertension was associated with male gender (PR = 1.54), lack of GP consultation (PR = 1.70) and being 25–44 years old (PR = 2.45) or 45–54 years old (PR = 2.37). Underreport of hypercholesterolemia was associated with lack of GP consultation (PR = 1.15), younger age (PR = 1.83 for 25–44 age group and PR = 1.52 for 45–54 age group), secondary (PR = 1.30) and higher (PR = 1.27) education. **Conclusion:** Self-reported data underestimate prevalence of hypertension and hypercholesterolemia. Magnitude of measurement error in self-reports varies by health conditions and population characteristics. Adding objective measurements to self-reported questionnaires improve data accuracy allowing better understanding of socioeconomic inequalities in health.

Introduction

Hypertension and hypercholesterolemia are major risk factors for cardiovascular diseases (CVDs) and contribute considerably to morbidity and mortality worldwide.[1–3] In Portugal, these conditions are recognized as priority for action. The National Priority Health Program on cerebro-cardiovascular diseases,[4] targeting the general population and specific social groups, has been dedicated to prevention and control of CVDs and its risk factors, with particular focus on hypertension and hypercholesterolemia.

High-quality epidemiological data are needed for monitoring CVD risk factors and planning public health interventions.[5] Many health indicators can be obtained through population-based surveys, using data from self-reports or objective measurements. Self-reports are frequently used to estimate the population-based prevalence of diseases and their determinants due to lower costs and faster data collection. However, quality of self-reported data has often been questioned due to bias.[6–8] Recall problems, undiagnosed diseases or misunderstanding of medical terms, are the major source of measurement error associated with self-reports.[6–9] Hypertension and hypercholesterolemia are chronic conditions known to have a long asymptomatic period, adding extra complexity to the use of the self-reported measures.

Although not consistent, there is some evidence in literature that measurement error in self-reported hypertension and hypercholesterolemia may differ between socioeconomic groups. Individuals with higher socioeconomic status (SES) may have better access to healthcare, better quality of healthcare and use it more frequently for preventive purposes,[10–12] and are therefore expected to have less undiagnosed diseases. In addition, more educated individuals have better understanding of health information and are more capable to answer survey questions on disease diagnosis.[11] As such, those with higher SES, are expected to report their health conditions more accurately in surveys. In China,[13] individuals with higher levels of income and education were more likely to report hypertension correctly. Low education was associated with less accurate report of hypercholesterolemia in Korea.[14] In Europe, no associations of reporting error and education were described by previous researches.[8,9]

If reporting errors occur more often in specific SES groups' inference about the health of the general population based on self-reported data on hypertension and hypercholesterolemia can be misleading with serious implications for public health monitoring and planning. Due to the high morbidity and mortality from cerebro-cardiovascular diseases in Portugal, it is crucial to evaluate to what extent objectively measured and self-reported health information differs in the Portuguese population.

This study aims to compare self-reported and examination-based prevalence estimates of hypertension and hypercholesterolemia and to identify factors associated with incorrect self-reports.

Methods

We used data from the first Portuguese National Health Examination Survey (INSEF), a cross-sectional population-based study based on a probabilistic sample ($n = 4911$) of non-institutionalized individuals aged between 25 and 74 years old, resident in Portugal for more than 12 months and able to understand an interview in Portuguese. The survey was carried out by the Instituto Nacional de Saúde Doutor Ricardo Jorge between 2013 and 2017 in collaboration with five Regional Health Administrations, the Regional Health and Social Affairs Offices of the Autonomous Regions of the Azores and Madeira, and with support from the Norwegian Institute of Public Health, through EEA Grants initiatives in public health.

The INSEF included blood pressure, weight, height, hip and waist circumferences measurements, and information on biochemical parameters determined on collected blood samples together with self-reported questionnaire data on socioeconomic and health status, health determinants and healthcare use. The fieldwork was conducted between February and December of 2015, using standardized measurement protocols of the European Health Examination Survey (EHES) and the experience gained from participating in the EHES Pilot Joint Action.[15,16] All interviews and examinations were conducted by health professionals at the local Health Centres. Health professionals involved in data collection had completed a specific training program on interview and examination protocols. The study was approved by the ethics committee of Instituto Nacional de Saúde Doutor Ricardo Jorge and Portuguese Data Protection Authority. Written informed consent was obtained from all participants. Detailed description of INSEF sample, survey procedures and characteristics of participants is provided elsewhere.[17]

Blood pressure was measured using an automated blood pressure measurement device (OMROM M6). All participants took five minutes rest in sitting position before blood pressure measurements. Three consecutive blood pressure measurements were taken on the right arm with a one minute interval between each measurement. For this study, the mean value of the 2nd and 3rd measurements was used.[18]

Non-fasting venous blood samples were collected and serum total cholesterol was measured through the cholesterol esterase/oxidase method. All laboratories involved and participated in the National Program for External Quality Assessment in 2015 to assure comparability and reliability of blood test results.

Participants' SES was described by education and income. We considered the highest level of education completed, grouped according the 2011 International Standard Classification of Education (ISCED)[19] into four categories: ISCED 0 level (no formal education/Basic(1 cycle)), ISCED 1–2 levels (Basic 2 cycle/ Basic 3 cycle), ISCED 3–4 levels (Secondary/Post-secondary), ISCED 5–8 levels (Higher/Post-graduate). Adult-equivalized available income was calculated using a modified OCDE equivalence scale[20] and categorized into five quintiles.

Self-reported prevalence of hypertension and hypercholesterolemia was based on the questions; 'Do you have any of the following diseases or conditions, diagnosed by a medical doctor? (Yes/No) High blood pressure or hypertension; Hypercholesterolemia.' For examination-based prevalence of hypertension and hypercholesterolemia, definitions proposed by the EHES guidelines[18] were used:

- Hypertension: the proportion of individuals with systolic blood pressure at least 140 mmHg or diastolic blood pressure at least 90 mmHg or who reported taking prescribed antihypertensive medication in two weeks prior the interview among all the survey participants.
- Hypercholesterolemia: the proportion of those who had a total serum cholesterol concentration of at least 190 mg/dl or reported using prescribed lipid-lowering medication in the two weeks prior to the interview among all the survey participants.

Examination-based and self-reported hypertension and hypercholesterolemia prevalence were estimated for the whole population and stratified by sex, age group, education, income, and consultation with general practitioner (GP) in 12 months prior to the interview. To assess the accuracy of self-reports of hypertension and hypercholesterolemia, sensitivity and specificity were calculated with the examination-based results, including current medication intake, as the reference standard. Following definitions were used:

- Sensitivity: the proportion of respondents who reported being diagnosed with hypertension/hypercholesterolemia among all those having hypertension/hypercholesterolemia according to the examination-based definitions.
- Specificity: the proportion of respondents who reported not having hypertension/hypercholesterolemia among all who do not have high blood pressure/high total serum cholesterol according to the examination-based definitions.

Design-adjusted Rao–Scott version of Pearson's chi-square test was used to compare sensitivity and specificity of self-reported hypertension and hypercholesterolemia among population subgroups.

Underreport of hypertension or hypercholesterolemia was defined as percentage of participants who fit the EHES hypertension and hypercholesterolemia definitions and did not report their conditions.

To identify factors independently associated with underreport of hypertension/hypercholesterolemia, prevalence ratios (PRs) were estimated using multivariate Poisson regression models adjusted for sex, age group, education, income and consultation with GP in previous years. Statistical analysis was conducted taking into account sampling weights that allowed obtaining prevalence estimates calibrated for the distribution of the Portuguese population by geographic region, age group, and sex. [SVY] package of Stata 11.2® software was used for data analysis. Significance level was set at 5%.

Results

Prevalence

Overall, 25.7% [23.9–27.5%] of the participants reported having hypertension and 24.8% [22.9–26.7%] hypercholesterolemia. The examination-based prevalence of hypertension and hypercholesterolemia was 36.0% [34.3–37.7%] and 63.3% [61.2–65.4%], respectively. Both, the self-reported and the examination-based prevalence of hypertension and hypercholesterolemia decreased with increase of educational level, while were less clearly patterned by income (table 1).

Self-reported data underestimated the prevalence for both health conditions in all population subgroups. For hypertension, self-reports led to underestimation of prevalence by 20.2% for women and by 36.6% for men, by 24.4% for the richest and by 31.7% for the poorest. The absolute difference between self-reported and examination-based prevalence decreased with increase of education from 13 percentage points (pp) in ISCED 0 group to 4.9 pp in ISCED 5–8 group, in contrast, relative difference was more pronounced among those more educated.

For hypercholesterolemia, self-reported data underestimated the prevalence by 59.1% for women and by 69.2% for men. Relative differences increased from 43.8% in ISCED 0 educational group to 77% in ISCED 3–4 group. Discrepancy between examination-based and self-reported hypercholesterolemia prevalence among income categories did not show a clear pattern, ranging between 60.7% and 62.2%.

Remarkable relative differences in prevalence rates were observed in young individuals (aged between 25–34 years old) for both hypertension (75.4%) and hypercholesterolemia (92.4%). In the 25–34 age group, the prevalence of hypertension and hypercholesterolemia according to self-reports was low. Given these data, the age groups 25–34 years and 34–44 years were aggregated in further analyses.

Sensitivity and specificity

Among the participants with hypertension according to examination-based data, 69.8% [64.8–74.4%] reported their health conditions correctly, while self-reporting sensitivity for hypercholesterolemia was only 38.2% [35.5–41.1%]. Statistically significant differences in sensitivity between population subgroups were verified for both health conditions (table 2). For hypertension, the sensitivity

Table 1 Self-reported and examination-based prevalence of hypertension and hypercholesterolemia

	Hypertension						**Hypercholesterolemia**					
	Examination-based		**Self-reported**		**Absolute difference**	**Relative difference (%)**	**Examination-based**		**Self-reported**		**Absolute difference**	**Relative difference (%)**
	%	**CI 95%**	**%**	**CI 95%**			**%**	**CI 95%**	**%**	**CI 95%**		
Overall	36.0	[34.3, 37.7]	25.7	[23.9, 27.5]	10.3	28.6	63.3	[61.2, 65.4]	24.8	[22.9, 26.7]	38.4	60.7
Sex												
Female	32.7	[30.1, 35.5]	26.1	[24.0, 28.3]	6.6	20.2	62.8	[60.3, 65.4]	25.7	[23.4, 28.1]	37.1	59.1
Male	39.6	[36.5, 42.8]	25.1	[22.1, 28.4]	14.5	36.6	63.8	[61.2, 66.4]	23.7	[21.3, 26.2]	40.1	62.9
Age group												
25–34 years	5.7	[3.1, 10.4]	1.4	[0.4, 4.5]	4.3	75.4	38.4	[32.9, 44.1]	2.9	[1.4, 5.9]	35.5	92.4
35–44 years	17.0	[14.6, 19.7]	9.6	[7.3, 12.4]	7.4	43.5	53.1	[50.1, 56.0]	9.7	[7.5, 12.4]	43.4	81.7
45–54 years	35.8	[31.3, 40.6]	19.8	[16.7, 23.3]	16.0	44.7	68.5	[65.0, 71.8]	22.3	[19.4, 25.4]	46.2	67.4
55–64 years	58.4	[51.4, 65.0]	47.1	[41.4, 53.0]	11.3	19.3	80.1	[77.6, 82.4]	44.7	[41.2, 48.3]	35.4	44.2
65–74 years	71.3	[65.7, 76.4]	58.9	[53.0, 64.6]	12.4	17.4	79.2	[75.2, 82.7]	50.5	[46.6, 54.4]	28.7	36.2
Stratum												
Rural	38.6	[34.5, 42.9]	26.7	[24.1, 29.6]	11.9	30.8	68.1	[65.5, 70.6]	26.6	[23.2, 30.3]	41.5	60.9
Urban	35.1	[33.3, 36.9]	25.3	[23.1, 27.6]	9.8	27.9	61.6	[59.0, 64.1]	24.1	[22.0, 26.4]	37.5	60.9
Education												
ISCED 0 level	62.6	[59.4, 65.8]	49.6	[47.0, 52.2]	13.0	20.8	76.4	[73.9, 78.7]	42.9	[39.3, 46.6]	33.5	43.8
ISCED 1–2 levels	33.4	[29.4, 37.6]	21.1	[17.3, 25.4]	12.3	36.8	64.0	[61.2, 66.8]	23.7	[20.8, 26.9]	40.3	63.0
ISCED 3–4 levels	24.0	[21.0, 27.4]	15.2	[12.2, 18.7]	8.8	36.7	57.5	[52.9, 61.9]	13.2	[10.1, 17.1]	44.3	77.0
ISCED 5–8 levels	15.5	[12.8, 18.5]	10.6	[7.9, 14.0]	4.9	31.6	49.8	[44.7, 54.9]	13.2	[10.7, 16.1]	36.6	73.5
Income												
1 quintile (lowest)	37.8	[32.6, 43.3]	25.8	[21.4, 30.7]	12.0	31.7	63.7	[59.7, 67.5]	24.1	[20.4, 28.2]	39.6	62.2
2 quintile	40.4	[35.7, 45.3]	29.9	[25.5, 34.8]	10.5	26.0	66.1	[61.3, 70.6]	26.0	[21.5, 30.9]	40.1	60.7
3 quintile	36.3	[32.7, 40.0]	25.3	[21.8, 29.1]	11.0	30.3	64.2	[60.3, 67.9]	24.9	[19.7, 31.0]	39.3	61.2
4 quintile	36.5	[32.2, 41.0]	25.0	[22.0, 28.3]	11.5	31.5	60.9	[56.8, 64.9]	23.4	[19.6, 27.7]	37.5	61.6
5 quintile (highest)	28.3	[25.3, 31.4]	21.4	[19.0, 24.0]	6.9	24.4	60.8	[55.3, 66.0]	22.6	[19.5, 26.1]	38.2	62.8
General practitioner consultation												
Less than 12 months	42.4	[39.7, 45.1]	33.3	[30.7, 35.9]	9.1	21.5	64.2	[62.0, 66.3]	29.3	[26.9, 31.8]	34.9	54.4
12 month or more	24.4	[20.8, 28.4]	12	[9.7, 14.9]	12.4	50.8	61.6	[58.2, 65.0]	16.3	[14.2, 18.6]	45.3	73.5

%: prevalence; CI 95%: 95% confidence interval.

Table 2 Sensitivity and specificity of the self-reported data on hypertension and hypercholesterolemia compared with examination-based data

	Hypertension				**Hypercholesterolemia**			
	Sensitivity		**Specificity**		**Sensitivity**		**Specificity**	
	%	**CI 95%**	**%**	**CI 95%**	**%**	**CI 95%**	**%**	**CI 95%**
Overall	69.8	[64.8, 74.4]	99.0	[98.5, 99.3]	38.2	[35.5, 41.1]	98.1	[96.8, 98.9]
Sex	***p* < 0.0001**		*p* = 0.2397		*p* = 0.1449		*p* = 0.3223	
Female	77.8	[71.9, 82.7]	98.8	[98.1, 99.2]	39.9	[36.2, 43.8]	97.7	[95.5, 98.8]
Male	62.5	[56.2, 68.3]	99.3	[98.4, 99.7]	36.4	[32.8, 40.1]	98.6	[97.0, 99.3]
Age group	***p* < 0.0001**		*p* = 0.2816		***p* < 0.0001**		***p* < 0.0001**	
25–44 years	51.5	[38.2, 64.7]	99.3	[98.5, 99.7]	13.7	[10.7, 17.4]	99.7	[98.7, 99.9]
45–54 years	53.5	[45.0, 61.9]	98.9	[97.7, 99.5]	32.4	[27.6, 37.5]	98.6	[95.7, 99.6]
55–64 years	79.9	[74.4, 84.5]	98.2	[95.2, 99.3]	54.9	[51.1, 58.6]	94.6	[88.3, 97.6]
65–74 years	82.2	[77.9, 85.8]	98.1	[93.3, 99.5]	62.1	[57.7, 66.3]	90.3	[82.6, 94.8]
Stratum	*p* = 0.6239		*p* = 0.8138		*p* = 0.9786		*p* = 0.8283	
Rural	68.3	[60.8, 74.9]	99.1	[98.1, 99.5]	38.3	[33.0, 43.9]	97.9	[94.0, 99.3]
Urban	70.4	[63.7, 76.3]	99.0	[98.4, 99.3]	38.2	[35.0, 41.5]	98.1	[96.7, 99.0]
Education	***p* = 0.0004**		***p* = 0.0319**		***p* < 0.0001**		***p* = 0.0009**	
ISCED 0 level	78.0	[74.1, 81.4]	97.2	[93.9, 98.8]	55.1	[50.9, 59.1]	94.3	[88.6, 97.2]
ISCED 1–2 levels	62.3	[53.9, 70.1]	99.6	[98.7, 99.9]	37.1	[32.8, 41.5]	99.4	[97.9, 99.8]
ISCED 3–4 levels	61.0	47.1, 73.4]	99.3	[97.6, 99.8]	21.7	[16.9, 27.4]	98.5	[95.8, 99.4]
ISCED 5–8 levels	63.9	[50.8, 75.2]	99.0	[97.3, 99.7]	24.7	[20.3, 29.7]	98.7	[96.2, 99.6]
Income	*p* = 0.296		*p* = 0.3792		*p* = 0.8232			
1 quintile (lowest)	67.5	[61.1, 73.2]	99.0	[97.0, 99.7]	36.9	[30.8, 43.4]	98.8	[96.0, 99.6]
2 quintile	73.8	[63.0, 82.3]	99.7	[99.5, 99.9]	38.7	[32.6, 45.3]	98.4	[93.6, 99.6]
3 quintile	67.0	[59.6, 73.6]	98.4	[96.6, 99.2]	39.4	[31.5, 48.0]	100.0	
4 quintile	67.2	[60.1, 73.5]	99.1	[96.7, 99.8]	37.1	[31.2, 43.4]	97.5	[94.2, 99.0]
5 quintile (highest)	73.3	[66.9, 78.8]	98.9	[97.2, 99.6]	35.3	[30.6, 40.4]	97.1	[93.5, 98.7]
General practitioner consultation	***p* < 0.0001**		*p* = 0.079		***p* < 0.0001**		***p* = 0.0081**	
Less than 12 months	77.0	[72.3, 81.1]	98.6	[97.7, 99.1]	44.7	[41.1, 48.4]	97.3	[95.6, 98.3]
12 month or more	48.2	[38.6, 57.8]	99.5	[98.7, 99.8]	25.8	[21.9, 30.1]	99.4	[97.9, 99.8]

%: percentage; CI 95%: 95% confidence interval.

P values of design-adjusted Rao–Scott version of Pearson's chi-square test, statistically significant values are in bold.

of self-reported data was higher among women 77.8% than men 62.5%. Sensitivity of self-reported hypertension and hypercholesterolemia was higher for participants of older age, and those who had consulted a GP in 12 months prior to interview (table 2). For both conditions sensitivity decreased with increase in educational attainments. For hypertension, the difference between the highest (ISCED 5–8) and the lowest (ISCED 0) level of education was modest (63.9% vs. 78.0%), while for hypercholesterolemia the difference was more profound (24.7% vs. 55.1%).

Specificity, i.e. the proportion of participants who do not have the condition and reported their status correctly, was high for both health conditions, 99% [98.5–99.3%] and 98.1% [96.8–98.9%] for hypertension and hypercholesterolemia, respectively. For hypertension, specificity varied marginally between age groups from 97.2–99.7%. For hypercholesterolemia, specificity decreased with age from 99.7% for 25–34 years old to 90.3% for 64–75 years old.

Factors associated with underreport of hypertension and hypercholesterolemia

Table 3 presents prevalence of underreport of hypertension and hypercholesterolemia, and adjusted prevalence ratios (PRs) of underreport according to sex, age group, education, income and consultation with GP in the last year.

Self-reported data failed to identify 30.2% of participants and considered as hypertensive according to the examination-based measurements. For hypercholesterolemia, the proportion of underreport was 61.8%.

Underreport of both hypertension and hypercholesterolemia was associated with younger age (25–44 years old and 45–55 years old groups) and lack of consultation with GP in recent years. Individuals from households in the 5th income quintile were less likely to misclassify themselves as not having hypertension or hypercholesterolemia (PR = 0.72 and PR = 0.87, respectively).

Men were more likely to underreport hypertension (PR = 1.54) than women. The effect of sex was not significant for underreport of hypercholesterolemia. For hypertension, the association of underreporting with educational level was not statistically significant, although reporting error seemed to be higher among more educated. For hypercholesterolemia, underreport was 30% higher (PR = 1.30) for individuals in ISCED 3–4 and 27% higher (PR = 1.27) for individuals in ISCED 5–8 education in comparison with those with low education (ISCED 0).

Discussion

This research, based on a large probabilistic sample representative for the Portuguese population aged 25–74 years, provides evidence on comparability of self-reported and objectively measured survey data on two major CVD risk factors. Our results showed that self-reported data underestimate the prevalence of hypertension and hypercholesterolemia.

For hypertension, the absolute difference between self-reported and examination-based prevalence was small (10.3 pp). For hypercholesterolemia, the difference was as high as 38.5 pp. This finding is in line with previous researches, where underestimation of hypercholesterolemia was verified for both men (49.6 pp) and women (43.6 pp).[7] Similar results (30.5-37.5pp) were reported for Spanish[9] and Australian populations,[6] however, it should be mentioned that in both studies a higher threshold was used to define hypercholesterolemia.

In self-report accuracy studies,[9,12,14] two types of measurement errors are distinguished. The first one is related with those who misclassified themselves as not having a disease, so-called underreport that corresponds to 1-sensitivity. The second one is overreport, related to those who do not have hypertension/hypercholesterolemia and misclassified themselves as having hypertension/hypercholesterolemia, that corresponds to 1-specificity.

In INSEF, overall hypertension sensitivity was 69.8%, varying across population subgroups from 48.2–82.2%. It means that about one-third of hypertension cases (31.2%) in the Portuguese population would be missed, if estimation of health indicators would rely solely on self-reported data. Compared with previous self-report validation studies from Europe[7–9] and other regions[6,13,14], these sensitivity values can be classified as moderate. Regarding previous researches in Portugal, sensitivity of hypertension self-report in INSEF was lower

Table 3 Prevalence of underreport of hypertension and hypercholesterolemia and adjusted prevalence ratios (PR) of underreport according to sex, age group, education, household equivalized disposable income quintile and consultation with GP

	Hypertension			Hypercholesterolemia		
	%	PR	CI 95%	%	PR	CI 95%
Overall	30.2			61.8		
Sex						
Female	22.2	1		60.1	1	
Male	37.5	1.54	[1.27, 1.86]	63.6	1.00	[0.92, 1.09]
Age group						
25–44 years	55.3	2.45	[1.94, 3.08]	86.3	1.83	[1.60, 2.10]
45–54 years	46.5	2.37	[1.73, 3.23]	67.6	1.52	[1.35, 1.72]
55–64 years	20.1	1.11	[0.79, 1.57]	45.1	1.08	[0.92, 1.26]
65–74 years	17.7	1		37.9	1	
Education						
ISCED 0 level	22.0	1		44.9	1	
ISCED 1–2 levels	37.7	1.10	[0.88, 1.38]	62.9	1.11	[1.00, 1.24]
ISCED 3–4 levels	39.0	1.11	[0.74, 1.69]	78.3	1.30	[1.15, 1.46]
ISCED 5–8 levels	36.1	1.08	[0.71, 1.62}	75.3	1.27	[1.12, 1.44]
General practitioner consultation						
Less than 12 months	23.0	1		55.3	1	
12 month or more	51.8	1.70	[1.37, 2.11]	74.2	1.15	[1.04, 1.26]
Income						
1 quintile (lowest)	32.5	1		63.1	1	
2 quintile	26.2	0.78	[0.56, 1.08]	61.3	0.94	[0.83, 1.06]
3 quintile	33.0	1.01	[0.83, 1.22]	60.6	0.88	[0.74, 1.05]
4 quintile	32.8	0.90	[0.73, 1.10]	62.9	0.91	[0.81, 1.02]
5 quintile (highest)	26.7	0.72	[0.56, 0.94]	64.7	0.87	[0.79, 0.95]

%: prevalence of underreport, PR: adjusted prevalence ratio, CI 95%: 95% confidence interval for adjusted prevalence ratio.

than in the 2011/12 PHYSA study (76.6%).[21] However, due to methodological differences of these studies, the comparisons should be interpreted with caution.

In terms of sensitivity of hypercholesterolemia, in INSEF only 38.2% reported their health status correctly, which is in line with other studies.[7,12,14] This sensitivity is considered low, indicating that two-thirds of all hypercholesterolemia cases (61.8%) are missed when estimation is based on self-reported data.

Specificity of self-report was high, above 98% for both conditions, meaning that participants who do not have hypertension/hypercholesterolemia rarely misclassified themselves as having hypertension/hypercholesterolemia. One possible explanation for this overreporting error could be that individuals earlier diagnosed were able to reduce their blood pressure/cholesterol to normal levels through lifestyle modification (diet or physical exercise) [22–24] and, therefore, did not fit our examination-based hypertension/hypercholesterolemia definition. Another possible explanation may be due to misreporting of cases whenever monitored blood pressure or cholesterol were very close to the 190 mg/dl threshold, but evidence was not enough for clinical diagnosis of the condition.[8,12]

Measurement error due to underreporting, i.e. participants who have hypertension or hypercholesterolemia and fail to report their health condition correctly, was more frequent than overreport. Among possible explanations for underreporting one might be undiagnosed, asymptomatic diseases in individuals who had no regular contact with GP.[8,11] Also, diagnosed individuals who were taking antihypertensive or lipid-lowering medication for some time before the interview may perceive their 'controlled' condition as not having diseases.[7] Our analysis showed that among participants who failed to report their health conditions in questionnaire, about 8.8% reported use of medication for lowering blood pressure levels and 2.9% for controling blood lipid levels. Thus, it seems that undiagnosed diseases constitute a major source of underreport. Both hypertension and hypercholesterolemia are asymptomatic, especially, in an early stage of disease.[7] Individuals, who were not in contact with healthcare system in 12 months prior to interview, had a lower probability of being diagnosed and, as such, to report their health status correctly.[8,11,13,25] In fact, we observed an association between the use of primary healthcare services and underreport, for both conditions.

Our results suggest that underreport error is unequally distributed among socioeconomic subgroups, leading to differential bias in prevalence estimates of self-reported health conditions for each group. Underreport of hypertension and hypercholesterolemia was associated with younger age and low income. Men were more likely to underreport hypertension and individuals with secondary (ISCED 3–4) or higher (ISCED 5–8) education were more likely to underreport hypercholesterolemia. Sex,[8,9] age[8,12] and education[13,14] are recognized in literature as important factors affecting surveys' response accuracy. Low sensitivity of self-reports among younger age groups for both hypertension[8,12] and hypercholesterolemia[6,9,12] was in accordance with previous studies. However, in our study, the effect of education on underreports of hypercholesterolemia was in the opposite direction than expected.[14] Although not statistically significant, increase of underreporting with increasing educational level was reported for the Australian population.[12] Also recently published NHANES data showed higher reporting error in hypercholesterolemia among college graduates.[25] In both studies, such findings were attributed to greater social desirability bias among more educated.[12,25] This may be true for Portuguese population, as well. Qualitative research showed that hypercholesterolemia perceived by patients as condition caused by unhealthy lifestyles, such as fatty food and lack of exercise,[26] while hypertension is more related to stress.[27] This may explain greater social desirability bias for hypercholesterolemia in relation to hypertension. In addition, it has been shown that hypercholesterolemia is not perceived by medical doctors as a disease.[28] Individuals with higher education are expected to have higher health literacy, more efficient GP-patient communication[11,28] and may, like GPs, perceive hypercholesterolemia as a risk factor and not as a disease. Hence, individuals with higher levels of education may not consider suitable to report it in survey when they are asked: 'Do you have any of the following diseases. . .?' So, careful wording of questions and their positioning in survey may also contribute to improve the accuracy of self-reported health conditions.

This study shares some limitation, common to the field of health examination survey research. First, examination-based prevalence of hypertension and hypercholesterolemia was defined based on measurements taken during a single health examination. In clinical practice, for a diagnosis, multiple measurements are usually required.[7,8,13] Second, participants' serum total cholesterol levels were measured on not-fasting blood samples. Also, the threshold of 190 mg/dl currently recommended in clinical practice is not consensual in literature.[29] These limitations could lead to bias towards overestimation of examination-based prevalence. Another methodological weakness is that INSEF participation rate was below 50%. However, the achieved participation rate of 43.9%[17] can be classified as average. It was comparable to participation rates of other health examination surveys recently developed in Europe, such as DEGS 2008-2011 and NLdeMaat 2010.[30] Also due to low prevalence of hypertension/hypercholesterolemia overreports, we were not able to perform multivariate analysis to identify participants' characteristics associated to overreporting.

Among relevant methodological strengths of our study are the use of standardized measurement protocols and continuous quality control[17] of all survey procedures.

In conclusion, our results suggest that for the Portuguese population aged 25–74 years, self-report underestimated the prevalence of hypertension and hypercholesterolemia. Specificity of self-report was high for both health conditions, while sensitivity was moderate for hypertension and low for hypercholesterolemia. Underreports were more frequent for hypercholesterolemia, than for hypertension.

Our study supports the hypothesis that measurement error in self-reports of hypertension and hypercholesterolemia depends on participants' characteristics. Use of self-reported data may lead to differential bias in prevalence estimates for certain population subgroups. Therefore, when estimating prevalence of hypertension and hypercholesterolemia or studying social determinants of these conditions with self-reported data, adjustments by sensitivity and specificity should be considered to account for differences in accuracy of self-reports among population sub-groups. Identification of factors associated underreport is important step for the development of statistical methods to minimize measurement error effect on parameters estimated in health surveys in data analysis phase.

Future public health policies and interventions should take into the considerations the socioeconomic pattern of underreported hypertension and hypercholesterolemia in Portugal.

Acknowledgements

The authors are grateful to all the professionals who were involved in the INSEF fieldwork and to all the INSEF participants and to Dr. Ricardo Mexia for English review.

Funding

The Portuguese National Health Examination Survey 2013-2017 (INSEF) was developed as part of the Pre-defined project of the Public Health Initiatives Program, 'Improvement of epidemiological health information to support public health decision and management in Portugal. Towards reduced inequalities, improved health, and bilateral cooperation', that benefits from a 1.500.000€

Grant from Iceland, Liechtenstein and Norway, through the EEA Grants.

Conflict of interest: None declared.

Key points

- Self-reported data underestimate the prevalence of hypertension and hypercholesterolemia in the Portuguese population.
- The measurement error in self-reports is unequally distributed among socioeconomic population sub-groups and varies with studied health conditions.
- Public health policies and interventions should take into consideration the socioeconomic pattern of underreported hypertension and hypercholesterolemia.
- The use of objective measures in population surveys can improve the accuracy of epidemiological information for monitoring CVD risk factors.

References

1 Farzadfar F, Finucane MM, Danaei G, et al. National, regional, and global trends in serum total cholesterol since 1980: systematic analysis of health examination surveys and epidemiological studies with 321 country-years and 3·0 million participants. *Lancet* 2011;377:578–86.

2 Forouzanfar MH, Afshin A, Alexander LT, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;388:1659–724.

3 Forouzanfar MH, Liu P, Roth GA, et al. Global burden of hypertension and systolic blood pressure of at least 110 to 115 mm Hg, 1990–2015. *JAMA* 2017;317:165.

4 Ministério da Saúde - Gabinete do MinistroDespacho n°16415/2003 (2.a Série). Diário da República n°193/2003, Série II de 2003-08-22, 12991–1299.

5 Direção Geral da S. *Programa Nacional Para as Doenças Cérebro-Cardiovasculares*. Portugal: Doenças Cérebro-Cardiovasculares em Números 2015. Lisboa, 2015.

6 Taylor A, Grande ED, Gill T, et al. Comparing self-reported and measured high blood pressure and high cholesterol status using data from a large representative cohort study. *Aust N Z J Public Health* 2010;34:394–400.

7 Tolonen H, Koponen P, Mindell JS, et al. Under-estimation of obesity, hypertension and high cholesterol by self-reported data: comparison of self-reported information and objective measures from health examination surveys. *Eur J Public Health* 2014;24:941–8.

8 Molenaar EA, Ameijden EJC, Van, Grobbee DE, Numans ME. Comparison of routine care self-reported and biometrical data on hypertension and diabetes: results of the Utrecht Health Project. *Eur J Public Health* 2007;17:199–205.

9 Huerta JM, José Tormo M, Egea-Caparrós JM, et al. Accuracy of self-reported diabetes, hypertension, and hyperlipidemia in the adult spanish population. DINO study findings. *Rev Esp Cardiol (Engl Ed)*. 2009;62:143–52.

10 Lorant V, Boland B, Humblet P, Delière D. Equity in prevention and health care. *J Epidemiol Community Health* 2002;56:510–6.

11 Burgard SA, Chen PV. Challenges of health measurement in studies of health disparities. *Soc Sci Med* 2014;106:143–50.

12 Peterson KL, Jacobs JP, Allender S, et al. Characterising the extent of misreporting of high blood pressure, high cholesterol, and diabetes using the Australian Health Survey. *BMC Public Health* 2016;16:1–11.

13 Ning M, Zhang Q, Yang M. Comparison of self-reported and biomedical data on hypertension and diabetes: findings from the China Health and Retirement Longitudinal Study (CHARLS). *BMJ Open* 2016;6:e009836.

14 Chun H, Kim IH, Min KD. Accuracy of self-reported hypertension, diabetes, and hypercholesterolemia: analysis of a representative sample of korean older adults. *Osong Public Health Res Perspect* 2016;7:108–15.

15 Tolonen H, Koponen P, Mindell J, et al. European Health Examination Survey-towards a sustainable monitoring system. *Eur J Public Health* 2014;24:338–44.

16 Kuulasmaa K, Tolonen H, Koponen P, et al. An overview of the European Health Examination Survey Pilot Joint Action. *Arch Public Health* 2012;70:20.

17 Instituto Nacional de Saúde Dr., Ricardo J. Inquérito Nacional de Saúde de Saúde com Exame Físico. Relatório metodológico [Portuguese Health Examination Survey: methodological report]. INSA.IP, editor. Lisboa, 2016.

18 Tolonen H. EHES manual part C. European level collaboration, 2013.

19 United Nations educational scientific and cultural organization. International standard classification of education, 2011.

20 Hagenaars AM, De Vos K, Zaidi A, editors. *Poverty statistics in the late 1980s: research based on micro-data*. Office for Official Publications of the European Communities. Luxembourg, 1994.

21 Polonia J, Martins L, Pinto F, Nazare J. Prevalence, awareness, treatment and control of hypertension and salt intake in Portugal. *J Hypertens* 2014;32:1211–21.

22 Diaz KM, Shimbo D. Physical activity and the prevention of hypertension. *Natl Inst Heal* 2013;15:659–68.

23 Toledo E, Hu FB, Estruch R, et al. Effect of the mediterranean diet on blood pressure in the PREDIMED trial: results from a randomized controlled trial. *BMC Med* 2013;11:207.

24 Mann S, Beedie C, Jimenez A. Differential effects of aerobic exercise, resistance training and combined exercise modalities on cholesterol and the lipid profile: review, synthesis and recommendations. *Sport Med* 2014;44:211–21.

25 Choi A, Cawley J. Health disparities across education: the role of differential reporting error. Health Econ [Internet]. 6 December 2018 [cited 2018 Mar 20]; 27. Available from: https://doi.org/10.1002/hec.3609.

26 Kirkegaard P, Edwards A, Risor MB, Thomsen JL. Risk of cardiovascular disease? A qualitative study of risk interpretation among patients with high cholesterol. *BMC Fam Pract* 2013;14:137.

27 Taylor C, Ward A. Patient's views of high blood pressure, its treatment and risks. *Aust Fam Physician* 2003;32:278–82.

28 Durack-Bown I, Giral P, D 'ivernois J-F, et al. Patient's and physician's perceptions and experience of hypercholesterolaemia: a qualitative study. *Br J Gen Pract* 2003;53:851–7.

29 Paalanen L, Koponen P, Laatikainen T, Tolonen H. Public health monitoring of hypertension, diabetes and elevated cholesterol: comparison of different data sources. *Eur J Public Health* 2018;28:754–65.

30 Mindell JS, Giampaoli S, Goesswald A, et al. Sample selection, recruitment and participation rates in health examination surveys in Europe–experience from seven national surveys. *BMC Med Res Methodol* 2015;15:78.

Supplementary Table Paper 1:

Distribution of Portuguese population and INSEF participants by sex and age group

	n	INSEF sample before weighting (%)	Population distribution (%)
Total	4911		
Sex			
Male	2265	46.1	47.5
Female	2646	53.9	52.5
Age group			
25-34	714	14.5	18.3
35-44	1135	23.1	23.5
45-54	1193	24.3	22.4
55-64	1098	22.4	19.9
65-74	771	15.7	15.9

5.2 Do self-reported data accurately measure health inequalities in risk factors for cardiovascular disease?

The second paper was published in 2019 in the International Journal of Public Health. In this paper I compared estimates of health inequality indicators based on the self-reported and objectively measured data and assessed effect of differential measurement error in outcome variable on estimates of absolute and relative inequalities by educational status.

International Journal of Public Health
https://doi.org/10.1007/s00038-019-01232-1

ORIGINAL ARTICLE

Do self-reported data accurately measure health inequalities in risk factors for cardiovascular disease?

Irina Kislaya[1,2] · Julian Perelman[2] · Hanna Tolonen[3] · Baltazar Nunes[1,2]

Received: 30 August 2018 / Revised: 4 March 2019 / Accepted: 7 March 2019

Abstract

Objectives This study aimed to compare the magnitude of educational inequalities in self-reported and examination-based hypertension and hypercholesterolemia and to assess the impact of self-reported measurement error on health inequality indicators.

Methods We used the Portuguese National Health Examination Survey data (n = 4911). The slope index of inequality (SII) and the relative index of inequality (RII) were used to determine the magnitude of absolute and relative education-related inequalities.

Results Among the 25–49-year-old (yo) men, absolute and relative inequalities were smaller for self-reported than for examination-based hypertension ($SII_{eb} = 0.18$ vs. $SII_{sr} = -0.001$, $p < 0.001$; $RII_{eb} = 1.99$ vs. $RII_{sr} = 0.86$, $p = 0.031$). For women, the relative inequalities were similar despite differences in self-reported and examination-based hypertension prevalence. For hypercholesterolemia, self-reported relative inequalities were larger than examination-based inequalities among the 50–74-yo men ($RII_{sr} = 2.28$ vs. $RII_{eb} = 1.21$, $p = 0.004$) and women ($RII_{sr} = 1.22$ vs. $RII_{eb} = 0.87$, $p = 0.045$), while no differences were observed among 25–49-yo.

Conclusions Self-reported data underestimated educational inequalities among 25–49-yo men and overestimated them in older individuals. Inequality indicators derived from self-report should be interpreted with caution, and examination-based values should be preferred, when available.

Keywords Health examination survey · Health inequalities · Hypercholesterolemia · Hypertension · Self-report

Introduction

Cardiovascular diseases (CVD) are the main cause for premature mortality in Europe, with a noteworthy social and economic impact (Wilkins et al. 2017). Hypertension and hypercholesterolemia are two modifiable CVD risk factors, annually responsible for more than 2/3 of deaths by CVD (Wilkins et al. 2017). Timely diagnosis, effective management, and prevention of these conditions are crucial for improvement in overall population health.

Socioeconomic gradients in CVD and its risk factors are well established: Individuals with lower socioeconomic status (SES) tend to have poorer health outcomes and face greater risk factors (Sommer et al. 2015; de Mestral and Stringhini 2017). Reduction in unfair differences between SES groups constitutes an important public health challenge in Europe (Official Journal of the European Union 2014). To monitor progress on this goal, health inequalities are most frequently measured based on self-reported data

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00038-019-01232-1) contains supplementary material, which is available to authorized users.

✉ Irina Kislaya
irina.kislaya@insa.min-saude.pt

1 Departamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge, IP, Lisbon, Portugal

2 Centro de Investigação em Saúde Pública, Escola Nacional de Saúde Pública, Universidade NOVA de Lisboa, Lisbon, Portugal

3 Department of Public Health Solutions, National Institute for Health and Welfare (THL), Helsinki, Finland

Published online: 25 March 2019

from population surveys. However, self-report, as a method of data collection, possesses some limitations. Individuals may report information incorrectly due to underdiagnoses, social desirability (i.e., people reporting what they expect to be the desirable behavior), or low literacy (i.e., misunderstanding of medical terms or confusion between cure and drug-related controlled disease). The misreport of disease is expected to be greater among low-SES people because of lower literacy and lower access to diagnosis, related to affordability or discrimination (Burgard and Chen 2014; Choi and Cawley 2018). Hence, the misreport may affect not only the prevalence estimates, but also the SES-related inequality measures.

It has been acknowledged in the literature that self-reported data do not always accurately measure the clinical diagnoses of CVD (Eliassen et al. 2016) or its risk factors (Newell et al. 1999; Mosca et al. 2013; Tolonen et al. 2014b; Paalanen et al. 2018). Namely, self-reported data tend to underestimate the prevalence of obesity, hypercholesterolemia, hypertension, and diabetes (Newell et al. 1999; Mosca et al. 2013; Tolonen et al. 2014a; Paalanen et al. 2018). Health examination surveys, which in addition to interview assemble biomarkers collection and physical examination, have been extensively used in the last two decades to perform studies on accuracy of self-reported data on CVD risk factors (Paalanen et al. 2018); however, to our best knowledge, little evidence is available regarding how the socioeconomic distribution of reporting errors affects health inequality estimates.

In the USA, according to NHANES data, self-reporting overestimated the educational disparities in hypercholesterolemia and underestimated disparities in hypertension and diabetes, when comparing prevalence rates among educational level groups (Choi and Cawley 2018). In the European context of universal health care, Mackenbach et al. (1996) showed that reporting error in diabetes and heart disease varied by level of education and self-reported data underestimated inequalities in the Netherlands. In Ireland, the educational gradient in hypertension was underestimated by self-reported data, while for hypercholesterolemia, no association with educational attainments was observed for either self-reported or objectively measured cholesterol (Mosca et al. 2013). Vellakkal et al. (2015) demonstrated, using a concentration index, that socioeconomic inequalities in hypertension in low- and middle-income countries may be underestimated or even have an opposite direction when self-reported data are used. However, none of those studies used regression-based inequality measures such as slope index of inequality (SII) or relative index of inequality (RII), which allow to account for the complete distribution of SES and not only extreme categories (Mackenbach and Kunst 1997; Speybroeck et al. 2012).

Although not consistent, these findings suggest that measurement error in self-reported CVD risk factors may differ between socioeconomic groups. In such a case, survey estimates based on self-reports may be more distant from "true" values for some population subgroups than for others, resulting in biased estimates of associations between health outcomes and their social determinants. So additional studies are needed to provide a more comprehensive view on the impact of measurement error from self-reported survey data on estimates of socioeconomic inequalities in CVD risk factors.

The study objectives were to: (1) compare the magnitude of educational inequalities in self-reported and examination-based hypertension and hypercholesterolemia in the Portuguese population and (2) assess the impact of self-reported measurement error on absolute and relative regression-based health inequality indicators (SII and RII).

Methods

Study design

We used data from the Portuguese National Health Examination Survey (INSEF), which has been described elsewhere (Nunes et al. 2018). Briefly, INSEF is a cross-sectional nationwide survey conducted in 2015 on a probability sample of community-dwelling individuals aged between 25 and 74 years old, resident in Portugal for more than 12 months, and able to follow an interview in Portuguese. The INSEF sample was designed to be representative at the national level as well as at the level of autonomous regions and the five mainland health regions. The sample was selected using a two-stage stratified probability-based cluster design.

INSEF combined information on measured biochemical parameters (total cholesterol, glycated hemoglobin, and blood count) and blood pressure measurements with a questionnaire applied through computer-assisted personal interview on demographic and socioeconomic characteristics, health conditions, medication intake, and health care use.

The fieldwork took place between February and December 2015 in primary care settings that offered all the necessary facilities for the survey implementation. Recruitment of participants was performed by 43 trained interviewers; interviews, physical examinations, and blood collections were conducted by 74 health professionals who had completed a 21-h training program on standardized survey procedures.

All measurements in INSEF were conducted using a standardized measurement protocol based on the recommendations of the Feasibility of a European Health

Examination Survey and the European Health Examination Survey Pilot Joint Action (Tolonen et al. 2008; Kuulasmaa et al. 2012) projects. All participants (n = 4911) provided written informed consent before data collection. INSEF was approved by the Ethics Commission of the National Health Institute Doctor Ricardo Jorge and by National Data Protection Authority (Authorization no. 9348/2010).

For each cardiovascular disease risk factor, analyses were limited to individuals with complete self-reported and examination-based data, and individuals with missing data (n = 25 for hypertension and n = 104 for hypercholesterolemia) were excluded from the analysis.

Definitions

Individuals were considered to have self-reported hypertension/hypercholesterolemia if they answered positively to both questions: "Do you have any of the following diseases or conditions: High blood pressure or hypertension; hypercholesterolemia? (Yes/No)" and if yes, "Were these conditions diagnosed by a medical doctor? (Yes/No)." Examination-based hypertension and hypercholesterolemia were based on objective measures of health conditions and self-reported use of corresponding medication. Information of medication intake was obtained from two questions: "During the past 2 weeks, have you used any medicines that were prescribed for you by a doctor?" and if yes, "Were the medicines for hypertension/hypercholesterolemia? (Yes/No)." Note that in the self-reported hypertension/hypercholesterolemia, the medication was not accounted for in the definition. This is because the Portuguese National Health Interview Survey, such as its European counterpart, does not include questions about specific medications for specific diseases. This failure is expected to contribute to the underestimation of prevalence and to the bias in inequality measurement.

Examination-based hypertension was defined as having: (1) systolic blood pressure of at least 140 mmHg, or (2) diastolic blood pressure of at least 90 mmHg, or (3) reported use of antihypertensive medication prescribed by a doctor, in 2 weeks prior to the interview.

Examination-based hypercholesterolemia was defined having total serum cholesterol concentration of at least 190 mg/dL or reported use of prescribed lipid-lowering medication in the 2 weeks prior to the interview.

The cutoff points for examination-based definitions were based on the current European and national clinical guidelines for CVD prevention (Reiner et al. 2011; Fifth Joint Task Force of the European Society of Cardiology et al. 2012; Direção Geral da Saúde 2013).

Participants' SES was measured through the highest level of education completed according to the 2011 International Standard Classification of Education (ISCED-2011) (United Nations Educational Scientific and Cultural Organization 2011). Four educational groups were considered: ISCED 0–1 levels (no formal education/basic [(1 cycle)/basic (2 cycles)]), ISCED 2 level [basic (3 cycles)], ISCED 3–4 levels (secondary/postsecondary), and ISCED 5–8 levels (higher/postgraduate).

Statistical analysis

Proportions of individuals with examination-based and self-reported hypertension and hypercholesterolemia were estimated at national level and stratified by sex, age group, and level of education.

To determine the magnitude of inequalities in CVD risk factors between the highest and lowest educational groups, we used the slope index of inequality (SII) and relative index of inequality (RII) (Mackenbach and Kunst 1997; Ernstsen et al. 2012). Mathematical formulation of RII and SII is described in detail elsewhere (Mackenbach and Kunst 1997; Ernstsen et al. 2012). SII and RII are regression-based inequality measures that take into account socioeconomic positions of population subgroups and their relative size. RII can be interpreted as a prevalence ratio between the most educated and the less educated, and SII represents the absolute difference in prevalence rates between the top and the bottom of educational hierarchy (Mackenbach and Kunst 1997).

To account for repeated measurements (self-reported and examination based) for the same individuals, we used generalized estimating equations (GEE). To estimate SII and RII, population was ranked from the highest (0) to lowest (1) level of education, and for each educational group, the *ridit score* was assigned based on midpoint of cumulative distribution of individuals. RII was estimated by log-link Poisson GEE with robust standard errors and an exchangeable working correlation structure, including age, *ridit*, type of measurement, and the *ridit* * type of measurement interaction. Statistically significant interaction term indicates that RIIs estimated with self-reported and examination-based data are different. Similar approach was used to assess inequalities in absolute scale. SII was estimated by Poisson GEE with identity link function. Poisson model was used since it provides direct estimates of prevalence ratios and differences and is recommended in the literature for cross-sectional studies (Barros and Hirakata 2003). Estimates of inequalities were stratified by age and sex, using four population groups [men 25–49-yo, women 25–49-yo, men 50–75-yo, and women 50–75-yo].

All statistical analyses were performed using sampling weights, to provide nationally representative results. The data were analyzed using the [SVY] package of Stata 15.1® software (StataCorp 2017). The significance level for all analyses was set at 5%.

Results

Sample characteristics

In total, of 4911 individuals participated in the INSEF survey (43.9% participation rate), 2265 (47.5%) were men and 2646 (52.5%) were women (Table 1). The majority had an ISCED 0–1 education level (40.3%). Hypertension was reported by 25.7% of participants, while 35.9% were considered to have hypertension according to examination-based data. Hypercholesterolemia was reported by 24.9%, but the proportion of individuals with measured high levels of total cholesterol was considerably higher (63.2%).

Prevalence

For hypertension, the self-reported and examination-based prevalence showed similar educational patterns in all four groups; namely, the proportion of hypertensive people was the highest among those with the lowest education levels (Fig. 1, Table S2). In general, the difference between self-reported and examination-based data was more pronounced among men than among women. The highest difference between self-reported and examination-based hypertension [23.5 percentage points (pp)] was observed among younger men with ISCED 2011 levels 0–1, i.e., the lowest level of education. The most accurate report of hypertension was verified for 25–49-yo women with ISCED 2011 levels 5–8 education.

For hypercholesterolemia, in general, the differences between self-reported and examination-based data were substantially larger than for hypertension (Fig. 2, Table S3). Likewise, for hypertension, the highest difference between proportion of self-reported and examination-based hypercholesterolemia was registered for younger men with the lowest level of education (48.7 pp).

Self-report of hypercholesterolemia varied considerably across educational categories in two population subgroups (25–49-yo women and 50–74-yo men). In both situations, individuals with the lowest level education had the highest prevalence. For examination-based hypercholesterolemia, the educational differences were less evident, in particular among older individuals. Among 50–74-yo women, the highest proportion of examination-based hypercholesterolemia (83.1%) was observed for those with the highest education levels.

Table 1 General participants' characteristics, Portuguese National Health Examination Survey 2015

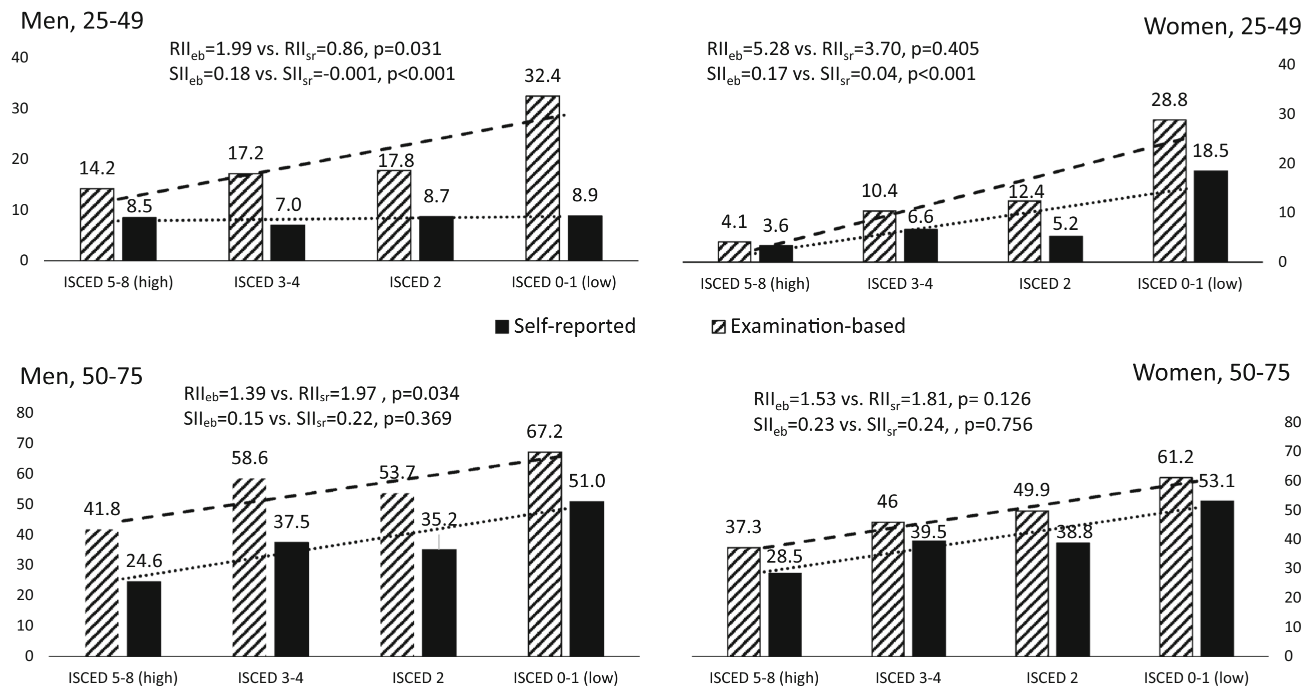
Participants' characteristics	n	%
Sex (n = 4911)		
Women	2646	52.5
Men	2265	47.5
Age group (n = 4911)		
25–49	2422	52.8
50–74	2489	47.2
Education (n = 4907)		
ISCED 2011 levels 0–1*	2193	40.3
ISCED 2011 level 2	918	18.9
ISCED 2011 levels 3–4	958	21.4
ISCED 2011 levels 5–8	838	19.4
Self-reported hypercholesterolemia (n = 4807)		
No	3573	75.1
Yes	1234	24.9
Examination-based hypercholesterolemia (n = 4807)		
No	1604	36.8
Yes	3203	63.2
Self-reported hypertension (n = 4886)		
No	3580	74.3
Yes	1306	25.7
Examination-based hypertension (n = 4886)		
No	3065	64.1
Yes	1821	35.9

**ISCED* International Standard Classification of Education 2011

Inequalities in hypertension

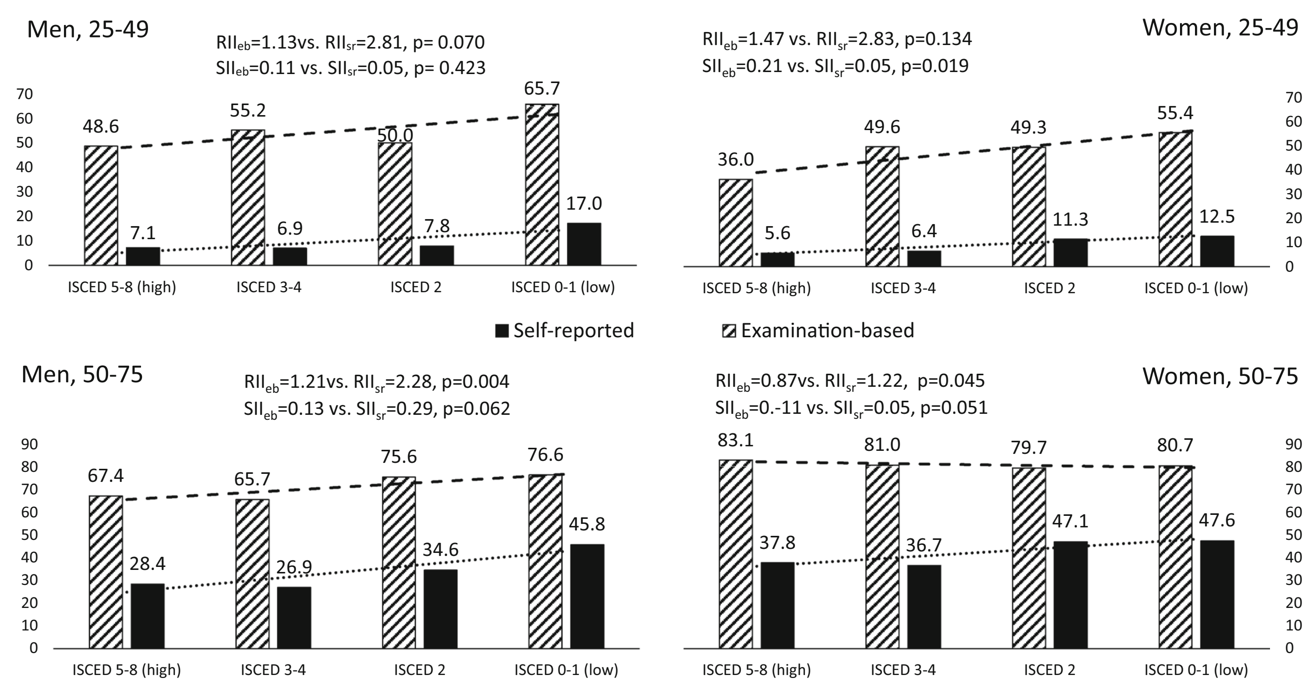
For survey results in general, relative inequalities were similar in magnitude for self-reported and examination-based (RII_{sr} = 1.74 vs. RII_{eb} = 1.76, p = 0.912) hypertension, indicating a lower prevalence of disease among the highly educated (Fig. 1, Table S4). Absolute inequalities were smaller for self-reported hypertension compared to examination-based hypertension ($p < 0.001$). Namely, according to self-reports, the discrepancy in hypertension prevalence between the lowest and highest educational categories was about 6% (SII_{sr} = 0.06), while according to examination-based data, the difference was 18% (SII_{eb} = 0.18).

Age- and sex-specific results showed considerable discrepancies in inequality indicators between self-reported and examination-based data. Among young men, self-reported data underestimated both absolute and relative inequalities. Namely, examination-based RII was 2.31 times as high as self-reported RII (RII_{eb} = 1.99 vs. RII_{sr} = 0.86, p = 0.031). Difference between examination-based and self-reported absolute inequalities estimates (18 pp) was statistically significant as well. (SII_{eb} = 0.18 vs. SII_{sr} = − 0.001, $p < 0.001$). Among 25–49-yo women, absolute inequalities were greater in magnitude for examination-based hypertension (SII_{eb} = 0.17 vs. SII_{sr} = 0.04, $p < 0.001$), while difference in relative inequalities estimates (RII_{eb} = 5.28 vs. RII_{sr} = 3.70, p = 0.405) was not statistically significant. Among 50–74-yo men, self-



* ISCED - International Standard Classification of Education 2011

Fig. 1 Proportion of participants with self-reported and examination-based hypertension according to educational level and absolute (SII) and relative (RII) inequality indexes stratified by age group and sex, Portuguese National Health Examination Survey 2015



* ISCED - International Standard Classification of Education 2011

Fig. 2 Proportion of participants with self-reported and examination-based hypercholesterolemia according to educational level stratified by age group and sex, and absolute (SII) and relative (RII) index of inequality, Portuguese National Health Examination Survey 2015

reported data overestimated relative educational inequalities (RII_{sr} = 1.97 vs. RII_{eb} = 1.39, p = 0.034). Difference in magnitude of self-reported and examination-based absolute inequalities in hypertension in this group was not statistically significant (SII_{sr} = 0.22 vs. SII_{eb} = 0.15, p = 0.369). Among 50–74-yo women, absolute and relative inequalities in self-reported and examination-based hypertension were similar ($SII_{sr} = 0.24$ vs. $SII_{eb} = 0.23$, $p = 0.756$; $RII_{sr} = 1.81$ vs. $RII_{eb} = 1.53$, $p = 0.126$).

Inequalities in hypercholesterolemia

For survey results in general, self-reported data overestimated educational gradient in hypercholesterolemia ($RII_{sr} = 1.67$ vs. $RII_{eb} = 1.12$, $p = 0.001$) (Fig. 2, Table S5). Although RII point estimates were greater in magnitude for self-reported data in all population subgroups, statistically significant gap in relative inequalities between examination-based and self-reported data was verified for 50–74-yo men ($RII_{sr} = 2.28$ vs. $RII_{eb} = 1.21$, $p = 0.004$) and women ($RII_{sr} = 1.22$ vs. $RII_{eb} = 0.87$, $p = 0.045$). Absolute inequalities in self-reported and examination-based hypercholesterolemia were similar for all population subgroups, except for 25–49-yo women. In this group, absolute inequalities were greater in magnitude for examination-based hypertension ($SII_{eb} = 0.21$ vs. $SII_{sr} = 0.05$, $p = 0.019$), suggesting a 21% difference in disease prevalence between women with the highest and lowest educational attainments, compared to 5% difference in prevalence obtained with self-reported data.

Discussion

Key findings

Using nationally representative data, this study examined how participants' self-reporting errors affected estimates of absolute and relative health inequalities for two major CVD risk factors: hypertension and hypercholesterolemia. In Portugal, self-reported data underestimated the prevalence of hypertension and hypercholesterolemia for the overall sample, and for all educational groups. This is in line with the previous research (Newell et al. 1999; Mosca et al. 2013). However, the magnitude of differences and, as such, the reporting error varied by educational attainment, suggesting a differential bias in education-related inequality indicators based on self-reported data. Among 25–49-yo men, the social gradients in self-reported hypertension systematically underestimated absolute and relative inequalities, while among 50–74-yo men, self-reported data overestimated relative inequalities, although to a smaller extent. Among women, self-reported data measured relative inequalities more accurately in both 25–49-yo and 50–74-yo groups. For hypercholesterolemia, self-reported data slightly overestimated inequalities. There was greater reporting bias among the highly educated 50–74-yo men and women.

Interpretation

In younger men, hypertension was more prevalent among those with lower education, although in a group more individuals failed to report it, leading to underestimation of absolute and relative inequalities. Less accurate report of hypertension by those with lower education levels in our study is in line with the previous research from the USA (Choi and Cawley 2018). Underreporting of disease in health surveys is attributed in the literature to underdiagnoses and lack of "awareness," which are strongly related to health literacy, health care access, health care quality, and type/form of health care use (Molenaar et al. 2007; Burgard and Chen 2014; Kulhánová et al. 2014; Tolonen et al. 2014b; Vellakkal et al. 2015). Hypertensive individuals, in particular at early stages of disease, may not experience any symptom (Tolonen et al. 2014b), so they may consider themselves healthy and do not seek medical care. However, men with higher education have a better understanding of the importance of disease prevention, so even without symptoms they are more likely to engage in screening programs and are more likely to be diagnosed (Lorant et al. 2002; Cutler and Lleras-Muney 2010; Kulhánová et al. 2014). Also due to higher health literacy and better access to health information, the highly educated have a greater ability to recognize symptoms, and therefore to discuss them with health professionals and be diagnosed (Burgard and Chen 2014). Besides universal health care coverage in Portugal, there still may exist some barriers in access to health care among young men with low socioeconomic status, explaining more undiagnosed diseases in this group. Other sex–age subgroups did not show the pattern observed for young men. In women, self-reported and examination-based inequality estimates were similar, while in older men, self-reported data overestimated inequalities in hypertension. Although these differences are difficult to explain, the gender gap in reporting errors in the young age group may be related to differences in health care use between men and women. Men use health care less frequently, are less likely to receive preventive care (Jeffries and Grogan 2012; Perelman et al. 2012), and consequently are less aware of their hypertension status and report it less accurately than women (Zhang and Moran 2017). This gender gap may be less present among older people, when both men and women are more likely to suffer from poor health and equally likely to use medical care (even, some studies show a lower use among women) (Cameron et al. 2010).

For hypercholesterolemia, use of self-reported data led to overestimation of educational gradient. This result reflects the underreporting of disease diagnosis by individuals with higher educational attainments in older age group. This may be explained by social desirability bias (Newell et al. 1999; Burgard and Chen 2014; Choi and Cawley 2018). It has been shown in the literature that individuals with higher education are more capable of identifying "sensitive" questions and are more concerned about their self-presentation (Preisendörfer and Wolter 2014), such that individuals may be more deceptive to report socially stigmatized diseases and unhealthy lifestyles. Similar results regarding the direction of educational gradient in reporting errors were found in the USA with NHANES data (Choi and Cawley 2018). Moreover, hypercholesterolemia is often linked to obesity, which is more prevalent among the worse-off (Gaio et al. 2017), so that medical doctors may be more prone to screen for cholesterol among the patients with low educational attainments because of their weight excess. Such discrimination in screening may also contribute to overestimation of educational inequalities in hypercholesterolemia. Lastly, more educated people may be more prone to consider hypercholesterolemia as a risk factor and not as a disease (Durack-Bown et al. 2003), and hence, they may not report it when interviewed about their health conditions.

Strengths and limitations

The INSEF survey presents methodological strengths: use of standardized measurement protocols, interviewers training, and continuous quality control of all survey procedures. Noticeably, it has the unique contribution in Portugal to include self-reported and examination-based ones for the same persons, which allows to compare indicators. The evaluation of inequalities was based on education as a measure of SES, which is stable along the life course, easy to report and collect, and less subject to reverse causality (von dem Knesebeck et al. 2006; Kulhánová et al. 2014; Campos-Matos et al. 2016; Choi and Cawley 2018). The use of ISCED 2011 contributes to comparability of our results at an international level. Finally, we reported both absolute and relative inequality measures, since both are recognized to be important for monitoring inequalities and public health planning (Mackenbach and Kunst 1997; Speybroeck et al. 2012).

Among limitations, the INSEF survey had only non-fasting samples to determine the lipid profile, and single physical examination as proxy for medical diagnosis of hypertension and hypercholesterolemia (Molenaar et al. 2007; Tolonen et al. 2014b). Also the 190 mg/dl cutoff currently recommended in clinical practice for definition of hypercholesterolemia is not consensual in the literature. However, a sensitivity analysis confirmed our findings when using an alternative cutoff value of 200 mg/dL (Table S6). Another methodological weakness is related to achieve participation rate of 43.9%, which can be considered average (Mindell et al. 2015). The INSEF participation rate varied with age, being the lowest among 25–34-yo (36%) and the highest among 55–64-yo (49.1%), which may also contribute for differences in inequality indicators between younger and older individuals. However, we should mention that the distribution of survey sample by sex and age was very close to population figures (Nunes et al. 2018).

Conclusions

Our results illustrated the significant impact of self-reported measurement error on estimates of socioeconomic inequalities in CVD risk factors. There is no straightforward universal answer about the direction and magnitude of the reporting bias. The survey response accuracy depends on the CVD risk factors, age, sex, and educational attainments of the respondents. The use of self-reported data may lead to underestimation of educational inequalities in some situations and overestimation in others. Inequality indicators derived from self-report should be interpreted with caution, and examination-based values should be preferred, when available.

Remarkable educational inequalities among young individuals raise important public health concerns regarding the increase in adverse CVD outcomes in the future, in particular among individuals with lower SES. These results also mean that there are opportunities for intervention to reduce health inequalities in CVD risk factors among the youngest. Development and implementation of specific preventive measures targeting younger age groups with low perception of being at risk of CVD may have important public health benefits.

Acknowledgements The Portuguese National Health Examination Survey (INSEF) was developed as part of the pre-defined project of the Public Health Initiatives Program, "Improvement of epidemiological health information to support public health decision and management in Portugal. Towards reduced inequalities, improved health, and bilateral cooperation," that benefits from a 1.500.000€ Grant from Iceland, Liechtenstein, and Norway through the EEA Grants and Portuguese Government. The authors are grateful to all professionals who were involved in the INSEF implementation and to all INSEF participants.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Barros AJ, Hirakata VN (2003) Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol 3:21. https://doi.org/10.1186/1471-2288-3-21

Burgard SA, Chen PV (2014) Challenges of health measurement in studies of health disparities. Soc Sci Med 106:143–150. https://doi.org/10.1016/j.socscimed.2014.01.045

Cameron KA, Song J, Manheim LM, Dunlop DD (2010) Gender disparities in health and healthcare use among older adults. J Women's Health 19:1643–1650. https://doi.org/10.1089/jwh.2009.1701

Campos-Matos I, Russo G, Perelman J (2016) Connecting the dots on health inequalities: a systematic review on the social determinants of health in Portugal. Int J Equity Health 15:1–10. https://doi.org/10.1186/s12939-016-0314-z

Choi A, Cawley J (2018) Health disparities across education: the role of differential reporting error. Health Econ. https://doi.org/10.1002/hec.3609

Cutler DM, Lleras-Muney A (2010) Understanding differences in health behaviors by education. J Health Econ 29:1–28. https://doi.org/10.1016/j.jhealeco.2009.10.003.Understanding

de Mestral C, Stringhini S (2017) Socioeconomic status and cardiovascular disease: an update. Curr Cardiol Rep 19:115

Direção Geral da Saúde (2013) Abordagem Terapêutica das Dislipidemias no Adulto. Lisboa

Durack-Bown I, Giral P, d'Ivernois J-F et al (2003) Patients' and physicians' perceptions and experience of hypercholesterolaemia: a qualitative study. Br J Gen Pract 53(496):851–857

Eliassen B-M, Melhus M, Tell GS et al (2016) Validity of self-reported myocardial infarction and stroke in regions with Sami and Norwegian populations: the SAMINOR 1 survey and the CVDNOR project. BMJ Open 6:e012717. https://doi.org/10.1136/bmjopen-2016-012717

Ernstsen L, Strand BH, Nilsen SM et al (2012) Trends in absolute and relative educational inequalities in four modifiable ischaemic heart disease risk factors: repeated cross-sectional surveys from the Nord-Trøndelag Health Study (HUNT) 1984–2008. BMC Public Health 12:266. https://doi.org/10.1186/1471-2458-12-266

Fifth Joint Task Force of the European Society of Cardiology, European Association of Echocardiography, European Association of Percutaneous Cardiovascular Interventions et al (2012) European guidelines on cardiovascular disease prevention in clinical practice (version 2012). Eur J Prev Cardiol 19:585–667. https://doi.org/10.1177/2047487312450228

Gaio V, Antunes L, Namorado S et al (2017) Prevalence of overweight and obesity in Portugal: results from the First Portuguese Health Examination Survey (INSEF 2015). Obes Res Clin Pract 12:40–50

Jeffries M, Grogan S (2012) 'Oh, I'm just, you know, a little bit weak because I'm going to the doctor's': young men's talk of self-referral to primary healthcare services. Psychol Health 27:898–915. https://doi.org/10.1080/08870446.2011.631542

Kulhánová I, Hoffmann R, Eikemo TA et al (2014) Educational inequalities in mortality by cause of death: first national data for the Netherlands. Int J Public Health 59:687–696. https://doi.org/10.1007/s00038-014-0576-4

Kuulasmaa K, Tolonen H, Koponen P et al (2012) An overview of the European health examination survey pilot joint action. Arch Public Health 70:20. https://doi.org/10.1186/0778-7367-70-20

Lorant V, Boland B, Humblet P, Deliège D (2002) Equity in prevention and health care. J Epidemiol Community Health 56:510–516. https://doi.org/10.1136/jech.56.7.510

Mackenbach JP, Kunst AE (1997) Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe. Soc Sci Med 44:757–771. https://doi.org/10.1016/S0277-9536(96)00073-1

Mackenbach JP, Looman CWN, Van Der Meer JBW (1996) Differences in the misreporting of chronic conditions, by level of education: the effect on inequalities in prevalence rates. Am J Public Health 86:706–711. https://doi.org/10.2105/AJPH.86.5.706

Mindell JS, Giampaoli S, Goesswald A et al (2015) Sample selection, recruitment and participation rates in health examination surveys in Europe—experience from seven national surveys. BMC Med Res Methodol 15:78. https://doi.org/10.1186/s12874-015-0072-4

Molenaar EA, Van Ameijden EJC, Grobbee DE, Numans ME (2007) Comparison of routine care self-reported and biometrical data on hypertension and diabetes: results of the Utrecht Health Project. Eur J Public Health 17:199–205. https://doi.org/10.1093/eurpub/ckl113

Mosca I, Bhuachalla BN, Kenny RA (2013) Explaining significant differences in subjective and objective measures of cardiovascular health: evidence for the socioeconomic gradient in a population-based study. BMC Cardiovasc Disord 13:64. https://doi.org/10.1186/1471-2261-13-64

Newell SA, Girgis A, Sanson-Fisher RW, Savolainen NJ (1999) The accuracy of self-reported health behaviors and risk factors relating to cancer and cardiovascular disease in the general population. Am J Prev Med 17:211–229. https://doi.org/10.1016/S0749-3797(99)00069-0

Nunes B, Barreto M, Gil AP et al (2018) The first Portuguese National Health Examination Survey (2015): design, planning and implementation. J Public Health (Bangkok). https://doi.org/10.1093/pubmed/fdy150

Official Journal of the European Union (2014) Regulation (EU) No 282/2014 of the European Parliament and of the Council of 11 March 2014 on the establishment of a third programme for the Union's action in the field of health (2014–2020). Off J Eur Union 13

Paalanen L, Koponen P, Laatikainen T, Tolonen H (2018) Public health monitoring of hypertension, diabetes and elevated cholesterol: comparison of different data sources. Eur J Public Health. https://doi.org/10.1093/eurpub/cky020

Perelman J, Fernandes A, Mateus C (2012) Gender disparities in health and healthcare: results from the Portuguese National Health Interview Survey. Cad Saude Publica 28:2339–2348. https://doi.org/10.1590/S0102-311X2012001400012

Preisendörfer P, Wolter F (2014) Who is telling the truth? A validation study on determinants of response behavior in surveys. Public Opin Q 78:126–146. https://doi.org/10.1093/poq/nft079

Reiner Z, Catapano AL, De Backer G et al (2011) ESC/EAS guidelines for the management of dyslipidaemias: the task force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS). Eur Heart J 32:1769–1818. https://doi.org/10.1093/eurheartj/ehr158

Sommer I, Griebler U, Mahlknecht P et al (2015) Socioeconomic inequalities in non-communicable diseases and their risk factors: an overview of systematic reviews. BMC Public Health 15:1–12. https://doi.org/10.1186/s12889-015-2227-y

Speybroeck N, Harper S, De Savigny D, Victora C (2012) Inequalities of health indicators for policy makers: six hints. Int J Public Health 57:855–858. https://doi.org/10.1007/s00038-012-0386-5

StataCorp (2017) Stata Statistical Software: Release 15. StataCorp, College Station

Tolonen H, Koponen P, Aromaa A et al (2008) Recommendations for the health examination surveys in Europe. Julkaisija-Utgivare-Publisher, Helsinki

Tolonen H, Koponen P, Mindell J et al (2014a) European health examination survey-towards a sustainable monitoring system. Eur J Public Health 24:338–344

Tolonen H, Koponen P, Mindell JS et al (2014b) Under-estimation of obesity, hypertension and high cholesterol by self-reported data: comparison of self-reported information and objective measures from health examination surveys. Eur J Public Health 24:941–948. https://doi.org/10.1093/eurpub/cku074

United Nations Educational Scientific and Cultural Organization (2011) International standard classification of education. United Nations Educational Scientific and Cultural Organization, Paris

Vellakkal S, Millett C, Basu S et al (2015) Are estimates of socioeconomic inequalities in chronic disease artefactually narrowed by self-reported measures of prevalence in low-income and middle-income countries? Findings from the WHO-SAGE survey. J Epidemiol Community Health 69:218–225. https://doi.org/10.1136/jech-2014-204621

von dem Knesebeck O, Verde PE, Dragano N (2006) Education and health in 22 European countries. Soc Sci Med 63:1344–1351. https://doi.org/10.1016/j.socscimed.2006.03.043

Wilkins E, Wilson L, Wickramasinghe K et al (2017) European cardiovascular disease statistics 2017 edition, vol 34. European Heart Network, Brussels, p 192. ISBN 978-2-9537898-1-2

Zhang Y, Moran AE (2017) Trends in the prevalence, awareness, treatment, and control of hypertension among young adults in the United States, 1999 to 2014. Hypertension 70:736–742. https://doi.org/10.1161/HYPERTENSIONAHA.117.09801

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

International Journal of Public Health

Do self-reported data accurately measure health inequalities in risk factors for cardiovascular disease?

Supplementary material

Table S1. Distribution of survey participants by education level, stratified by age group and sex, Portuguese National Health Examination Survey 2015

		ISCED 2011*			
		Level 0-1	Level 2	Level 3-4	Level 5-8
Men	25-49	22.6	26.6	28.3	22.5
	50-74	63.7	15.0	14.2	7.1
Women	25-49	17.8	17.9	30.9	33.4
	50-74	62.0	15.9	10.2	11.9

* ISCED - International Standard Classification of Education 2011

Table S2. Proportion of participants with self-reported and examination-based hypertension according educational level, stratified by age group and sex, Portuguese National Health Examination Survey 2015

			ISCED 2011*			
			Level 0-1	Level 2	Level 3-4	Level 5-8
Men	25-49 years (n=1,091)	Examination-based	32.4% [26.5,38.8]	17.8% [11.5,26.6]	17.2% [11.9,24.1]	14.2% [8.2, 23.6]
		Self-reported	8.9% [4.0,18.7]	8.7% [4.6,16.1]	7.0% [3.4,13.8]	8.5 % [3.7,18.4]
		Absolute difference	-23.5[-31.9, -15.1]	-9.1 [-15.6, -2.5]	-10.2 [-17.1, -3.2]	-5.7 [-9.8, -1.5]
	50-75 years (n=1,161)	Examination-based	67.2% [60.5,73.2]	53.7% [41.5, 65.5]	58.6% [46.9, 69.4]	41.8% [24.6, 61.2]
		Self-reported	51.0% [46.6, 55.3]	35.2% [25.0, 47.0]	37.5% [25.5, 51.3]	24.6% [12.4, 43.1]
		Absolute difference	-16.2 [-21.7, -10.7]	-18.5 [-25.0, -12.0]	-21.1[-30.9, -11.2]	-17.2 [-31.3, -2.9]
Women	25-49 years (n= 1,325)	Examination-based	28.8% [20.0, 39.5]	12.4% [6.8, 21.4]	10.4% [5.6,18.5]	4.1% [2.4, 6.9]
		Self-reported	18.5% [11.3, 28.7]	5.2% [2.2, 11.5]	6.6% [3.5,12.1]	3.6% [1.5, 8.2]
		Absolute difference	-10.3 [-15.0, -5.6]	-7.2 [-12.3, -2.2]	-3.8 [-7.7, 0.01]	-0.5 [-3.6, 2.5]
	50-75 years (n=1,305)	Examination-based	61.2% [55.9, 66.2]	49.9% [39.9, 60.0]	46.0% [32.9, 59.7]	37.3 % [27.2, 48.6]
		Self-reported	53.1% [48.1, 58.1]	38.8% [28.0, 50.8]	39.5% [26.1, 54.6]	28.5% [19.7, 39.2]

		Absolute difference	-8.1[-10.9, -5.2]	-11.1 [-20.7, -1.6]	-6.5[-12.2, -0.9]	-8.8 [-15.8, -1.7]
* ISCED - International Standard Classification of Education 2011						

Table S3. Proportion of participants with self-reported and examination-based hypercholesterolemia (cut-off 190 mg/dL) according educational level, stratified by age group and sex, Portuguese National Health Examination Survey 2015

			ISCED 2011*			
			Level 0-1	Level 2	Level 3-4	Level 5-8
Men	25-49 years (n=1079)	Examination-based	65.7 % [53.6,76.1]	50.0 % [40.9,59.1]	55.2 % [46.0,64.0]	48.6% [38.4,58.9]
		Self-reported	17.0% [11.7,24.1]	7.8% [5.1,11.8]	6.9 % [3.2,14.1]	7.1% [3.0,16.0]
		Absolute difference	-48.7 [-58.4, -39.0]	-42.2[-51.2, -33.2]	-48.3[-57.2, -39.4]	-41.5[-48.4, -34.6]
	50-75 years (n=1147)	Examination-based	76.6% [71.8,80.8]	75.6% [66.2,83.0]	65.7% [53.0,76.5]	67.4% [50.7,80.6]
		Self-reported	45.8% [40.4,51.2]	34.6 % [22.7,48.8]	26.9% [18.4,37.6]	28.4% [17.9,41.9]
		Absolute difference	-30.8[-37.0, -24.7]	-41.0[-55.2, -26.8]	-38.8[-48.2, -29.4]	-39.0[-58.7, -19.3]
Women	25-49 years (n=1307)	Examination-based	55.4% [43.6,66.6]	49.3% [40.4,58.3]	49.6% [44.3,54.8]	36.0 % [28.7,44.1]
		Self-reported	12.5% [7.5,20.0]	11.3% [7.0,17.9]	6.4% [2.9,13.3]	5.6 % [3.2,9.6]
		Absolute difference	-42.9 [-52.2, 33.7]	-38.0[-48.6, -27.5]	-43.2 [-51.1, -35.3]	-30.4 [-39.1, -21.7]
	50-75 years (n=1274)	Examination-based	80.7% [76.7,84.2]	79.7 % [69.0,87.4]	81.0% [69.0,89.1]	83.1% [72.7,90.1]
		Self-reported	47.6% [42.5,52.8]	47.1% [34.1,60.5]	36.7% [23.4,52.5]	37.8 % [28.6,48.0]

		Absolute difference	-33.1[-39.2, -26.8]	-32.6[-42.0, -23.3]	-44.3[-55.3, -33.1]	-45.3[-60.0, -30.6]
* ISCED - International Standard Classification of Education 2011						

Table S4. Absolute (SII) and relative (RII) index of inequality for self-reported and examination-based hypertension, Portuguese National Health Examination Survey 2015

		Examination-based hypertension		Self-reported hypertension		RII_{eb} / RII_{sr} [95%CI]
	n	RII	CI 95%	RII	CI 95%	
Overall	4,882	**1.76**	[1.43, 2.17]	**1.74**	[1.35, 2.24]	1.01 [0.84, 1.22]
Men 25-49	1,091	**1.99**	[1.06, 3.71]	0.86	[0.32, 2.30]	**2.31** [1.08, 4.96]
Men 50-74	1,161	**1.39**	[1.05, 1.86]	**1.97**	[1.31, 2.97]	**0.71** [0.52, 0.98]
Women 25-49	1,325	**5.28**	[2.33, 11.9]	**3.70**	[1.26, 10.80]	1.42 [0.62, 3.31]
Women 50-74	1,305	**1.53**	[1.11, 2.11]	**1.81**	[1.24, 2.65]	0.85 [0.68, 1.04]
		SII	CI 95%	SII	CI 95%	SII_{eb} - SII_{sr} [95%CI]
Overall	4,882	**0.18**	[0.12, 0.24]	**0.06**	[0.03, 0.14]	**0.13** [0.08, 0.17]
Men 25-49	1,091	**0.18**	[0.07, 0.29]	-0.001	[-0.08, 0.08]	**0.18** [0.10, 0.27]
Men 50-74	1,161	0.15	[-0.1, 0.32]	**0.22**	[0.06, 0.37]	-0.07 [-0.21, 0.08]
Women 25-49	1,325	**0.17**	[0.11, 0.23]	0.04	[-0.02, 0.11]	**0.13** [0.07, 0.19]
Women 50-74	1,305	**0.23**	[0.07, 0.38]	**0.24**	[0.09, 0.39]	-0.02 [-0.11, 0.08]

Values in bold represent statistically significant results (2-sided P<0.05)

Table S5. Absolute (SII) and relative (RII) index of inequality for self-reported and examination-based hypercholesterolemia (cut-off 190 mg/dL) Portuguese National Health Examination Survey 2015

		Examination-based hypercholesterolemia		Self-reported hypercholesterolemia		RII_{eb} / RII_{sr} [95%CI]
	n	**RII**	CI 95%	**RII**	CI 95%	
Overall	4,807	1.12	[0.99, 1.26]	**1.67**	[1.29, 2.16]	**0.66** [0.52, 0.85]
Men 25-49	1,079	1.13	[0.86, 1.47]	**2.81**	[1.01, 7.91]	0.40 [0.15, 1.08]
Men 50-74	1,147	1.21	[0.97, 1.50]	**2.28**	[1.43, 3.64]	**0.53** [0.34, 0.81]
Women 25-49	1,307	**1.47**	[1.09, 2.00]	**2.83**	[1.13, 7.12]	0.52 [0.22, 1.22]
Women 50-74	1,274	0.87	[0.75, 1.01]	1.22	[0.86, 1.74]	**0.71** [0.51, 0.99]
		SII	CI 95%	SII	CI 95%	SII_{eb} - SII_{sr} [95%CI]
Overall	4,807	0.07	[-0.01, 0.14]	0.07	[0.01, 0.11]	0.001 [-0.07, 0.07]
Men 25-49	1,079	0.11	[-0.03, 0.26]	0.05	[-0.05, 0.14]	0.07 [-0.08, 0.22]
Men 50-74	1,147	0.13	[-0.03, 0.28]	**0.29**	[0.13, 0.44]	-0.16 [-0.33, 0.01]
Women 25-49	1,307	**0.21**	[0.07, 0.34]	0.05	[-0.01, 0.11]	**0.16** [0.03, 0.28]
Women 50-74	1,274	-0.11	[-0.23, 0,02][	0.05	[-0.10, 0.19]	-0.16 [-0.32, 0.001]

Values in bold represent statistically significant results (2-sided $P<0.05$)

Table S6. Absolute (SII) and relative (RII) index of inequality for self-reported and examination-based hypercholesterolemia (alternative definition, cut-off 200 mg/dL) Portuguese National Health Examination Survey 2015

		Examination-based hypercholesterolemia		RII_{eb} / RII_{sr} [95%CI]
	n	**RII**	CI 95%	
Overall	4,807	**1.18**	[1.03, 1.37]	**0.69** [0.55, 0.88]
Men 25-49	1,079	**1.40**	[1.02, 1.92]	0.51 [0.19, 1.32]
Men 50-74	1,147	**1.37**	[1.05, 1,79]	**0.60** [0.40, 0.90]
Women 25-49	1,307	1.43	[0.98, 2.08]]	0.52 [0.22, 1.19]
Women 50-74	1,274	0.86	[0.71, 1.04]	**0.72** [0.52, 0.99]
		SII	CI 95%	SII_{eb} - SII_{sr} [95%CI]
Overall	4,807	**0.09**	[0.2, 0.16]	0.03 [-0.03, 0.10]
Men 25-49	1,079	**0.21**	[0.06, 0.35]	**0.16** [0.02, 030]
Men 50-74	1,147	**0.19**	[0.03, 0.35]]	-0.10 [-0.26, 0.06]
Women 25-49	1,307	**0.15**	[0.03, 0.27]	**0.10** [0.02, 0.22]
Women 50-74	1,274	-0.09	[-0.23, 0.05]	-0.13 [-0.29, 0.02]

Values in bold represent statistically significant results (2-sided $P<0.05$)

5.3 Collecting valid and reliable data: fieldwork monitoring strategies in a Health Examination Survey

The third paper was published in 2020 in the Portuguese Journal of Public Health. It addresses the research question on how measurement error can be controlled during survey implementation. Fieldwork monitoring strategies aimed on quality assurance for self-reported and objectively measures are described as well as field work interventions.

Portuguese Journal of Public Health

Research Article

Port J Public Health 2020;38:81–90
DOI: 10.1159/000511576

Received: March 15, 2020
Accepted: September 5, 2020
Published online: December 14, 2020

Collecting Valid and Reliable Data: Fieldwork Monitoring Strategies in a Health Examination Survey

Irina Kislaya[a, b] Ana João Santos[a, b] Heidi Lyshol[c] Liliana Antunes[a] Marta Barreto[a, b] Vânia Gaio[a, b] Ana Paula Gil[d] Sónia Namorado[a, b] Carlos Matias Dias[a, b] Hanna Tolonen[e] Baltazar Nunes[a, b]

[a]Departamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisbon, Portugal; [b]Centro de Investigação em Saúde Pública, Escola Nacional de Saúde Pública, Universidade Nova de Lisboa, Lisbon, Portugal; [c]Department of Health and Inequality, Division of Mental and Physical Health, Norwegian Institute of Public Health, Oslo, Norway; [d]Interdisciplinary Centre of Social Sciences (NOVA FCSH), Lisbon, Portugal; [e]Department of Public Health Solutions, Finnish Institute for Health and Welfare (THL), Helsinki, Finland

Keywords
Health information · Data quality · Health examination survey · Fieldwork monitoring · Interviewer observation

Abstract
Introduction: Health surveys constitute a relevant information source to access the population's health status. Given that survey errors can significantly influence estimates and invalidate study findings, it is crucial that the fieldwork progress is closely monitored to ensure data quality. The objective of this study was to describe the fieldwork monitoring conducted during the first Portuguese National Health Examination Survey (INSEF) regarding protocol deviations and key performance indicators (KPI). ***Methods:*** Data derived from interviewer observation and from the statistical quality control of selected KPI were used to monitor the four components of the INSEF survey (recruitment, physical examination, blood collection and health questionnaire). Survey KPI included response rate, average time distribution for procedures, distribution of the last digit in a specific measure, proportion of haemolysed blood samples and missing values. ***Results:*** Interviewer observation identified deviations from the established protocols, which were promptly corrected. During fieldwork monitoring through KPI, upon implementation of corrective measures, the participation rate increased 2.5-fold, and a 4.4-fold decrease in non-adherence to standardized survey procedures was observed in the average time distribution for blood pressure measurement. The proportion of measurements with the terminal digit of 0 or 5 decreased to 19.6 and 16.5%, respectively, after the pilot study. The proportion of haemolysed samples was at baseline level, below 2.5%. Missing data issues were minimized by promptly communicating them to the interviewer, who could recontact the participant and fill in the missing information. ***Discussion/Conclusion:*** Although the majority of the deviations from the established protocol occurred during the first weeks of the fieldwork, our results emphasize the importance of continuous monitoring of survey KPI to ensure data quality throughout the survey.

I. Kislaya and A.J. Santos made equal contributions to this study and are co-first authors.

karger@karger.com
www.karger.com/pjp

Karger
OPEN ACCESS

Irina Kislaya
Departamento de Epidemiologia
Instituto Nacional de Saúde Doutor Ricardo Jorge
Avenida Padre Cruz, PT–1649-016 Lisbon (Portugal)
irina.kislaya@insa.min-saude.pt

Recolha de informação válida e credível em saúde: estratégias de monitorização do trabalho de campo num inquérito com exame físico

Palavras Chave

Informação em saúde · Qualidade dos dados · Inquérito de saúde com exame físico · Monitorização da recolha de dados · Observação do entrevistador

Resumo

Introdução: Os inquéritos de saúde constituem uma importante fonte de informação para conhecer o estado de saúde da população. Visto que os erros associados aos inquéritos podem afetar significativamente as estimativas, invalidando as suas conclusões, é crucial monitorizar o progresso do trabalho de campo. Este estudo teve como objetivo descrever a monitorização do trabalho de campo realizado durante o primeiro Inquérito Nacional de Saúde com Exame Físico (INSEF) referente a desvios ao protocolo e principais indicadores de desempenho (KPI). ***Métodos:*** Dados resultantes da observação dos entrevistadores e do controlo estatístico de qualidade de alguns dos KPI foram utilizados para monitorizar as quatro componentes do inquérito (recrutamento, exame físico; colheita de sangue e questionário de saúde), durante a implementação do trabalho de campo. Os KPI selecionados incluíram a taxa de resposta, distribuição do tempo médio de realização dos procedimentos, proporção do último dígito para medidas específicas, proporção de amostras de sangue hemolisadas e dos valores omissos. ***Resultados:*** A observação dos entrevistadores permitiu identificar e corrigir atempadamente desvios ao protocolo. Após a implementação de medidas corretivas, com base na monitorização dos KPI, a taxa de participação aumentou 2,5 vezes e foi observada uma redução de 4,4 vezes na não adesão aos procedimentos padronizados para a medição da pressão arterial. Após o estudo piloto, a proporção de medições com o dígito terminal de 0 ou 5 diminuiu para 19,6% e 16,5%, respectivamente. A proporção de amostras hemolisadas foi inferior a 2,5%. A proporção dos valores omissos foi minimizada comunicando-os imediatamente ao entrevistador, que poderia recontactar o participante e completar a informação. ***Discussão/Conclusão:*** Embora a maioria dos desvios ao protocolo tenha ocorrido durante as primeiras semanas do trabalho de campo, os resultados mostram a importância da sua monitorização continua nos inquéritos de saúde de forma a garantir a qualidade dos dados recolhidos.

Introduction

Knowledge of the populations' health is an essential step into making improvements that prompt healthy behaviours, identify and address adverse health events, and prevent and treat diseases [1]. Health surveys constitute one of several different information sources that allow this knowledge to be gathered. Health surveys represent important tools for public health planning and interventions, essential for the development and monitoring of national and regional health plans [2].

In the context of the health surveys, health examination surveys, where the information collected via a detailed questionnaire is complemented with objective information measured by physical examination and laboratory tests on biological samples, may provide more accurate and better-quality information [3]. However, the collection of objective information does not ensure data validity and reliability. The quality of a survey process, including its instruments and fieldwork adherence to study protocols, is crucial for accurate, reliable and valid results [4]. Considering that public health decisions are also based on health survey data, it is important to implement quality assurance procedures in order to prevent unacceptable practices and to minimize errors in data collection [4]. Survey errors can significantly influence estimates and invalidate study findings [5].

The total survey error approach [6] comprehensively conceptualizes the errors associated with the different survey production processes: the design, collection, processing and analysis of survey data [7]. Considering survey error as the deviation of a survey response from its underlying true value [7], total survey error identifies two major error components that influence data accuracy: representativeness and measurement errors [5].

Errors in surveys resulting from the misrepresentation of the target population, also denominated as "non-observation errors," usually include coverage error (discrepancies in the survey statistics between the target population and the frame population), sampling error and unit non-response [6]. Measurement errors are generally also referred to as "observational errors" [8] and include errors deriving from the interviewer, the respondent, the data collection mode, the information system and the interview setting [9].

Close fieldwork monitoring and tracking of field progress can help identify key indicators of non-response and measurement error, which can be actively used to improve data collection and, consequently, data quality [10]. Standardization of survey procedures through the detailed fieldwork protocol, survey staff training and continuous monitoring of overall survey performance through tangible performance indicators (key performance indicators [KPI]), calculated from paradata collected during the implementation of the survey, constitute the main strategies for minimizing measurement errors in health surveys.

This article aims to describe the KPI for fieldwork monitoring used in the first Portuguese National Health Examination Survey (INSEF) to evaluate overall survey performance and detect deviations from the survey protocol. It also describes the corrective measures implemented during the fieldwork to achieve the established performance targets.

Subjects and Methods

Survey Settings

INSEF was a cross-sectional survey aimed at collecting objective and self-reported data on health status, health determinants and use of health care in a representative probabilistic sample (n = 4,911) of Portuguese residents aged between 25 and 74 years [2]. The INSEF sample was selected using two-stage probabilistic cluster sampling. In the first stage, 7 geographical areas (primary sampling units [PSU]) were randomly selected in each of the 7 Portuguese health regions, and, afterwards, individuals were selected in each of the selected PSU.

The fieldwork was performed between February and December 2015 by 24 teams, which included 117 professionals. Each team comprised 1 administrative staff, 2 nurses (1 appointed as a team coordinator) and 1 laboratory technician or nurse to perform the blood collection and processing [11, 12]. All fieldwork staff underwent a training program of 21–28 h including role-play of all data collection procedures, and received four operating manuals, with step-by-step descriptions of standard operational procedures for each survey component. In each of the 49 PSU, data collection took place in primary care health centres for approximately 2 consecutive weeks.

The survey included four components: recruitment (R), physical examination (PE), venous blood collection (BC) and a structured health questionnaire (HQ) [2].

Recruitment

About 2 weeks prior to data collection, an invitation was sent by regular mail to the selected individuals with a signed letter from their general practitioner and the INSEF coordinator. For each PSU, in the week prior to data collection, the administrative staff contacted the selected individuals by telephone to schedule the interview and health examination.

Records of the recruitment attempts (up to 6) were kept for each of the selected individuals, including the date, time, outcome of each recruitment attempt, verification of eligibility criteria and reasons for refusal for non-participants. A participation rate of 40% was established as a target [12].

Physical Examination

PE comprised measurement of blood pressure (BP), height, weight, and waist and hip circumference. Measurements were performed according the European Health Examination Survey (EHES) guidelines [13].

The BP measurements were taken in the sitting position on the right arm. Three sequential readings were taken for each participant in an unhurried way [13], including 5 min of rest before the first measurement and a 1-min interval between measurements. Measurements taken over a period of less than 8 min (possibly without including rest) were considered non-compliant with the survey procedure.

For anthropometric measurements, the participants were asked to take off heavy clothes and empty their pockets. Weight was measured in kilograms to the nearest 0.1 kg, while weight and waist and hip circumference were measured to the nearest millimetre. Rounding of measured values was considered as a deviation from the measurement protocol.

Blood Collection

Venous blood samples were collected for determination of the lipid profile from all participants except those with anaemia or other chronic illnesses, which restricts taking blood samples. Up to 2 attempts of blood withdrawal were made. The serum samples were centrifuged 30–60 min after collection and transported daily to the participating laboratories to be fully processed.

Health Questionnaire

The HQ included 23 thematic sections, covering sociodemographic information, disease and chronic conditions, functional limitations, mental health, health determinants and healthcare use. It was applied by trained nurses by Computer-Assisted Personal Interviewing (CAPI) using the REDCap web application [14], while paper-based forms designed for optical recognition were used for the remaining three components (R, PE and BC) (see the INSEF fieldwork flowchart in online suppl. Fig. 1A; for all online suppl. material, see www.karger.com/doi/10.1159/000511576).

Fieldwork Monitoring Strategies

INSEF fieldwork monitoring strategies were organized in groups according to selected classes of the two types of survey error (Table 1). The strategies employed within the recruitment component aimed specifically at reducing non-response error, by focusing on gaining respondent cooperation and on a correct assessment of eligibility criteria. A separate set of monitoring strategies (interviewer observation and statistical quality control) focused on reduction of measurement error for R, PE, BC and HQ.

Interviewer Observation

INSEF planning led to the implementation of different types of interviewer observations throughout the course of the fieldwork for all survey components. A standardized observation grid was used for fieldwork personnel evaluation, and results were shared with the fieldwork team in an end-of-day interactive meeting between the supervisor and the team members and through a small written report.

The first observation took place during the 1-day pilot study conducted prior to the start of actual fieldwork, which included between 3 and 8 participants out of 20 selected individuals per region. Whenever possible, the supervisors would also accompany

Table 1. Fieldwork monitoring strategies by survey component (INSEF 2015)

Fieldwork monitoring strategies		Period/frame	Survey component(s)
Recruitment and participation control schedule	Scheduled interviews, no shows, rescheduled and completed interviews	Daily/weekly	R
Interviewer observation	Pilot	1 per region	R, PE, BC, HQ
	Accompanying of the fieldwork team in a new PSU	18	PE, BC, HQ
	Audit visit to the PSU (scheduled and non-scheduled)	20	R, PE, BC, HQ
	External quality assessment	3	PE, BC, HQ
	Telephone "mystery client"	42	R
	Face-to-face "mystery client"	4	PE, BC, HQ
Statistical quality control	Response rate	By observation site	R
	Item missing data	Weekly or twice a month (R, PE, BC); daily (HQ)	R, PE, BC, HQ
	Assessment of inconsistent data and outliers	Weekly or twice a month (PE + BC); daily (HQ)	PE, BC, HQ
	Assessment of time stamps	Twice a month (PE + BC); weekly (HQ)	PE, BC, HQ
	Proportion of haemolysed blood samples		BC
	Blood sample transport temperature variation		BC
	Verification of calibration values for measurement devices	Daily, weekly and twice a month (according to equipment)	PE, BC
	Verification of last digit distribution (physical examination procedures)	Pilot, observation sites and randomly during fieldwork for some interviewers	PE, BC
	Inter-observer variability (equal measurements)	Pilot and observation sites	PE, BC

PSU, primary sampling unit; R, recruitment; PE, physical examination; BC, blood collection; HQ, health questionnaire.

the fieldwork team during their first day in a new PSU. This activity took place, particularly, if a new team began collecting data for the first time or if we anticipated that a specific PSU could be particularly challenging due to logistical constraints or less engaged populations. The supervisors made several scheduled and non-scheduled audit visits to monitor data collection for 2 or more participants. The face-to-face observations were made with the survey participants' consent.

Finally, the "mystery client" observation procedure was used. In this procedure, one of the supervisors (unknown to the specific team and familiar with all survey procedures) pretended to be a survey participant in order to perform fieldwork personnel evaluation without revealing his/her true identity. A telephone mystery client was used for the recruitment stage, while a face-to-face mystery client was used for other survey components when the monitoring activities indicated a possible non-compliance with data collection procedures.

Statistical Quality Control and KPI

Statistical quality control was used to understand if a variation in any component of the survey process was natural and expected or a problematic process variation that needed to be addressed. It aimed to assure the proper application of interviewing and measuring protocols and hence to reduce measurement error.

KPI were chosen to be monitored daily, weekly or monthly to evaluate adherence to protocols based on EHES guidelines [13, 15], World Health Organization MONICA Project recommendations [16] and previous practical experience gathered from participation in the EHES pilot in 2009–2010 [17]. The periodicity of assessment varied by survey component; while the HQ web application allowed real-time monitoring and validation, R, PE and BC were verified with a 2-week delay in paper forms.

The 6 KPI monitored during the fieldwork (Table 2) included participation rate, average time distribution for procedures, proportion of a last digit in a specific measure and missing values. To monitor BC and processing, we selected the proportion of haemolysed samples as the main quality indicator. Sample haemolysis, defined as the rupturing of red blood cells, is one of the most important causes of pre-analytical errors that may compromise laboratory test results. It can occur due to incorrect procedures related to BC, namely due to use of a tourniquet for more than 1 min, or improper handling or storage of the samples [18].

The results describe the KPI ordered along the timeline of the fieldwork period. Each of the 7 health regions conducted the fieldwork in 7 primary care health centres (PSU). The results of all regions were combined and presented by sets of 7 observations throughout the survey period to evaluate trends within the survey – each group is composed of 1 PSU from each of the 7 health regions, corresponding to the different health regions first, second, third, fourth, fifth, sixth and seventh PSU of each.

Results

Recruitment

The recruitment KPI throughout the data collection period are displayed in Figure 1. After the pilot, we observed a 1.3- and 2.5-fold increase in contact and participation rates, respectively. The cooperation rate improved from 29.2% to 53.5% in the first group of PSU and remained above 64% afterwards. The average survey contact, cooperation and participation rates were 69.5%, 63.1% and 43.9%, respectively.

In the first days of fieldwork, the daily contact with the survey staff responsible for recruitment highlighted that, of all scheduled appointments, about 84% resulted in participation – some individuals who agreed to participate simply forgot the appointment (no show) – and that taking time off from work to participate in the survey was one of the main reasons for refusal. The "mystery client" procedure allowed us to detect some deviations from the recruitment protocol in eligible criterion verification and survey presentation.

Physical Examination

The distribution of time needed for three subsequent BP measurements is presented in Figure 2. The average time for BP measurement was 10 min. The proportion of measurements that lasted less than 8 min went down from 14.3% in the second PSU group to 3.3% in the seventh PSU group, indicating a 4.4-fold decrease in non-adherence to the standardized survey procedure.

For the measurements of height and waist and hip circumference, the protocol stated that the interviewer should not round up or down the value of the measurement. Figure 3 presents the distribution of last digits for height, which was expected to be relatively uniform. In the pilot, there was a clear preference for the digit 0

Table 2. Key performance indicators for fieldwork monitoring (INSEF 2015)

Survey component	Key performance indicators	Units	Target
Recruitment	Contact, cooperation and participation rate	Pilot and PSU for all the regions (PSU1 through PSU7)	40% participation rate
Physical examination	Average time for blood pressure measurement	Pilot and PSU for all the regions (PSU1 through PSU7)	Minimum duration of 8 min
	Verification of the last digit distribution for height	Pilot and PSU for all the regions (PSU1 through PSU7)	–
Blood collection and processing	Proportion of haemolysed serum samples	PSU for all the regions (PSU1 through PSU7)	Below 5%
Health questionnaire	Average interview time	PSUs for all the regions (PSU1 through PSU7)	–
	Average rate of items missing for selected questionnaire areas	Overall	Below 5%

PSU, primary sampling unit.

Port J Public Health 2020;38:81–90
DOI: 10.1159/000511576

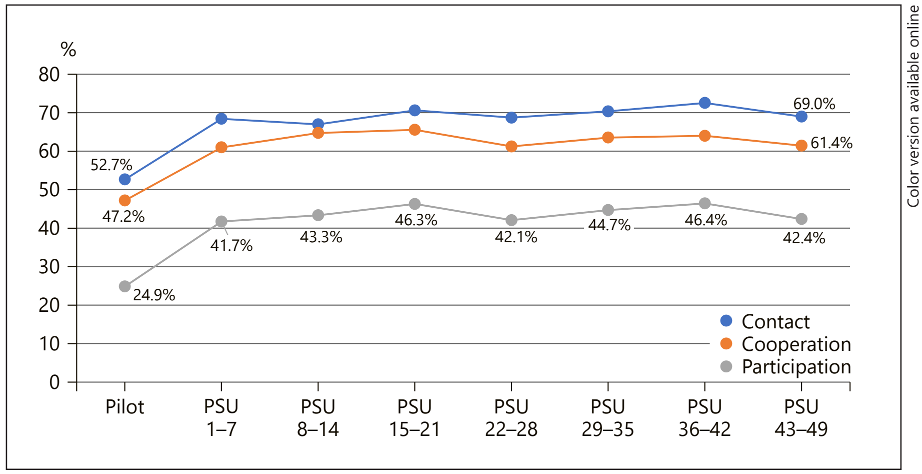


Fig. 1. Contact, cooperation and participation rates throughout the data collection period. PSU, primary sampling unit.

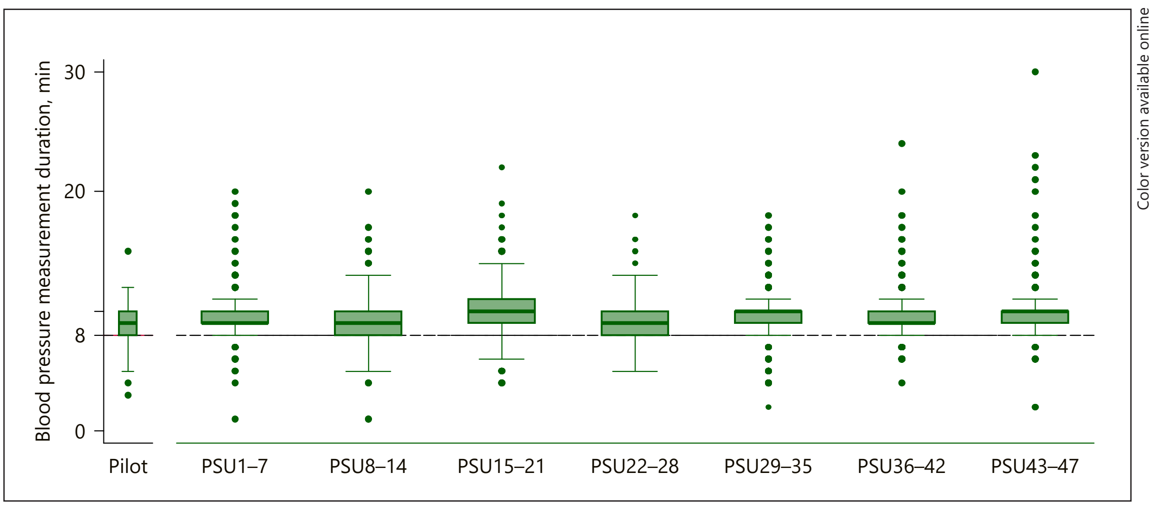


Fig. 2. Distribution of blood pressure measurement duration (min) over fieldwork. PSU, primary sampling unit.

(29.2%) and the digit 5 (24.4%), indicating that the interviewers were rounding up or down their values. The proportion of measurements with the terminal digit of 0 or 5 decreased to 19.6% and 16.5%, respectively, after the pilot, and remained below 20% afterwards, although with some variation between PSU groups.

Interviewer observation identified some deviations from the measurement protocol, such as incorrect participant posture during BP and anthropometric measurements.

Blood Collection

Blood samples were successfully collected from 98.8% of the participants. The proportion of haemolysed samples was below 2.5% in all PSU, ranging from 0.4% to 2.2% by PSU group (Table 3). Direct observation of BC identified some deviations from the protocol, including prolonged use of the tourniquet and centrifugation of blood tubes not within the recommended time, between 30 and 60 min.

Health Questionnaire

The HQ was administered by CAPI after PE and BC. The median interview time decreased over the fieldwork, from 34 to 28 min (Fig. 4). Considering as "missing" responses both "refused to answer" and "too uncertain to answer," the missing values varied from 0% to 14.6% (Table 4). For the core survey questions, the proportion of missing values was below 5%. The questions regarding

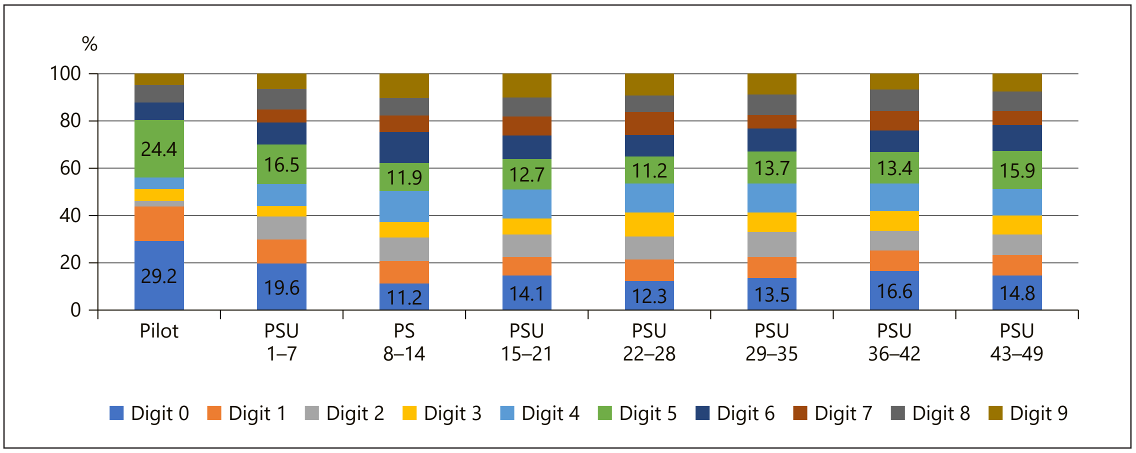


Fig. 3. Last digit distribution for height measurement throughout the fieldwork. PSU, primary sampling unit.

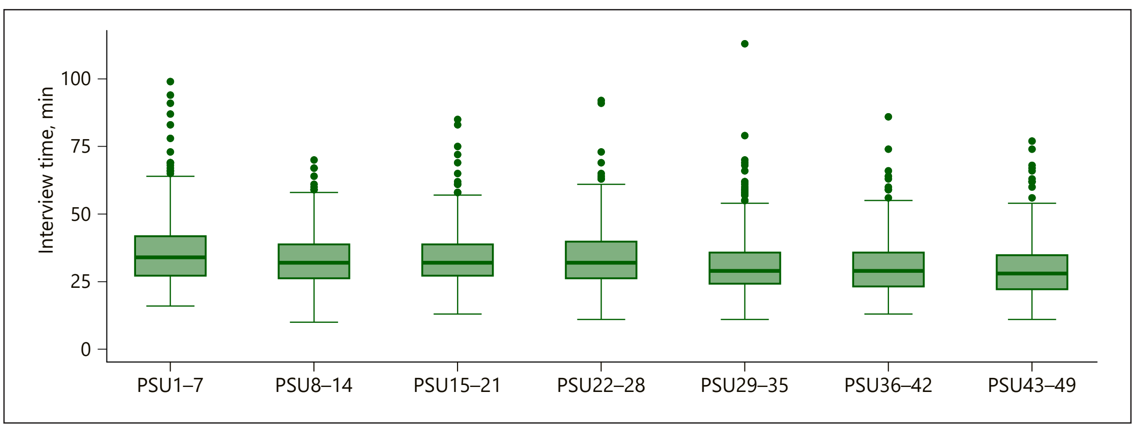


Fig. 4. Average interview time (min) throughout the fieldwork. PSU, primary sampling unit.

waiting time or duration presented higher proportions of missing values, up to 14.6%.

The observation of HQ administration also identified critical areas where interviewers made errors, for instance, by not presenting a response card or not reading the response options when required, or by adopting an inappropriate interviewing style.

Interventions to Improve Data Quality

The supervisors developed several actions when deviations from expected values for a specific KPI were observed or when fieldwork personnel observation identified non-adherence to the protocol. To guarantee the required number of participants within the established fieldwork plan, staff was instructed to "overbook" participants based on detected no-show rates, reducing the interval between appointments from 45 to 30 min. Additionally, to reduce no shows, a reminder call to the scheduled participant on the day before the appointment was added to the recruitment procedure. The data collection schedule was adjusted to include non-working hours in order to increase participation. Deviations from the protocol detected by the "mystery clients" were corrected through additional training sessions; 33 participants who failed to fulfil the eligibility criteria were excluded from the survey database in the data cleaning stage. Information on the number of participants at regional and national levels was compiled weekly into a newsletter called "INSEF barometer," which was shared with all the personnel involved in data collection and their supervisors.

Table 3. Proportion of haemolysed samples throughout the fieldwork in INSEF 2015

PSU1–7	PSU8–14	PSU15–21	PSU22–28	PSU29–35	PSU36–42	PSU43–49
1.8%	1.2%	0.8%	2.2%	0.8%	0.4%	1.3%

PSU, primary sampling unit.

Table 4. Missing item response values throughout the health questionnaire (INSEF 2015)

Question (all primary sampling units)	Refused to answer, %	Uncertain, %
Do you have any of the following diseases or conditions: high blood pressure or hypertension?	0.02	0.30
Do you have any of the following diseases or conditions: hypercholesterolemia?	0.04	0.97
When was the last time that your blood pressure was measured by a health professional?	0.02	2.73
When was the last time that you have made the following tests: cholesterol?	0.02	3.44
When was the last time that you have made the following tests: faecal occult blood test?	0.20	14.40
On a typical summer day, how much time do you usually spend outdoors between 10 a.m. and 4 p.m. in your leisure time (weekend, day off, holidays)?	0.12	2.00
Regarding your last GP visit, how many days did you wait between making an appointment and the consultation?	0.57	12.07

When PE, BC or HQ indicators suggested systematic errors, an email was sent to all fieldwork personnel, highlighting the need to correctly address non-adherence to the protocol (e.g. rounding up/down weight values of tenths to 0 or 5), as well as to the supervisors, contacting the fieldwork personnel by telephone or scheduling a site visit to provide additional training in person.

Data-related issues such as missing data and inconsistent item responses in the HQ data were also fed back almost immediately (on the same or the next day) to the interviewer, who could, if necessary, contact participants to clarify the registered answer. To improve interviewers' performance in the HQ component, written reminders and validation rules were added to some questions in the REDCap web application.

In addition, the results of statistical quality control were presented to all regional teams, and, if necessary, incorrect or inconsistent data were noted and the correct procedure was discussed.

Discussion and Conclusion

INSEF was implemented with a high concern about data quality during all its phases, combining quantitative and qualitative approaches. In addition to interviewer training and piloting of all survey procedures, we adopted a set of KPI for the data collected during R, PE, BC and HC. Recruitment monitoring highlighted the need to implement alternatives to meet the minimum necessary number of participants. Timely adjustments to the recruitment protocol ("overbooking," reminder calls and data collection schedule adjustments) resulted in considerable improvements after the pilot, although participation rates varied between the PSU groups. The changes in participation rate were not merely the result of varying experience among the personnel making the phone calls, but also that some PSU presented greater challenges. Overall, the achieved participation rate of 43.9% was lower than those reported for surveys based solely on interviewing (76–80%) [19], but higher than the rate previously obtained in the Portuguese EHES pilot in 2010 (36.8%) [17] and above the established target of 40% [12]. Compared to other health examination surveys in Europe, the INSEF participation rate was similar to that in a German survey (41–44%) and higher than that in a Scottish survey (36–41%), but lower than those in Finnish (56–61%) and Italian (54%) surveys [20].

INSEF was able to collect complete and valid information on the PE, BC and HQ components. The small losses in BC observed (1.2%) were below the 5% threshold defined by the MONICA project for the "completeness" data quality dimension [16]. The low proportion (<5%) of missing item response values for the core HQ areas may be a reflection of the validation routines at the time of data entry, which were included in the REDCap web application.

Standardization of the measurement procedures and quality assurance of the PE data were the most challenging. Regarding BP measurement, for more than 90% of the individuals observed, the suggested protocol was followed, with the total measurement time being longer than 8 min. Some deviance in the distribution of the last digit of the height measurement was observed in the pilot and the first PSU group, but afterwards the distribution became more uniform. However, the observed variation in terminal digit distribution suggests that even though the indicator stabilized over time, it was not completely corrected, and some interviewers kept rounding up and down values. Several retraining sessions were required to improve standardization of the height measurement procedure.

As expected, we observed considerable improvements for all KPI after the pilot study. Piloting was useful to verify the feasibility of the data collection procedures, adherence to the protocol and possible deviations, as well as to identify survey procedures that require closer monitoring throughout the fieldwork. For some KPI, the trend over the fieldwork period was not linear but rather represented some trendless fluctuation. These results illustrate a need for continuous monitoring of survey procedures and collection of paradata during the fieldwork.

In all interviewer-mediated surveys, interviewers play a crucial role during the entire data collection process [8, 9, 21]. Proper interviewer training and interviewer observation have been recognized in the literature as a key to successful survey implementation [9, 16]. INSEF quality control programs have benefited from the communication and relationship established between supervisors and fieldwork staff, as well as from the daily or weekly contact between supervisors and fieldwork staff. From training onwards, the supervisors have conveyed the importance of the standardized procedures and monitoring activities, including collection of paradata. A survey organization culture that emphasizes the importance of data quality and can make fieldwork staff understand that they are an integral part of achieving high-quality data is likely to lead to improved quality [5]. On the other hand, close monitoring and the availability of supervisors to provide assistance encouraged the fieldwork staff to talk about any situations they encountered in the field that were not covered in the training, or any remaining doubts or questions. Overall, by checking the fieldwork staff's work regularly and providing positive feedback, the supervisors could ensure that the quality of data collection remained high throughout the survey.

Among the limitations of our quality assurance approach it should be mentioned that in INSEF several corrective actions were taken simultaneously to improve data quality and measurement protocol adherence, which makes it difficult to identify the individual impact of any specific corrective action on either sample representativeness or adherence to the interview protocols.

The use of paradata and monitoring of the survey process offer the opportunity to make real-time decisions informed by observations of the ongoing data collection process. Because the production of survey data involves many actors, the importance of high data quality must be made a priority for all those involved, and the interviewers' role in this process should be seen as that of collaborators.

Acknowledgements

The authors are grateful to all the professionals who were involved in the implementation of INSEF and to all INSEF participants.

Statement of Ethics

Ethical approval by the National Health Institute Ethics Committee and National Data Protection Commission was obtained.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

INSEF was developed as part of the predefined project of the Public Health Initiatives Program, “Improvement of Epidemiological Health Information to Support Public Health Decision and Management in Portugal. Towards Reduced Inequalities, Improved Health, and Bilateral Cooperation,” that benefited from a EUR 1,500,000 grant from Iceland, Liechtenstein and Norway through the EEA Grants and the Portuguese government.

Author Contributions

I. Kislaya and A.J. Santos contributed substantially to the conception and design of the study; the acquisition, analysis and interpretation of the data; and drafting of the first version of the manuscript. H. Tolonen and B. Nunes contributed substantially to the conception and design of the study and to data interpretation and provided a critical revision of the article. H. Lyshol, L. Antunes, M. Barreto, V. Gaio, A.P. Gil, S. Namorado and C.M. Dias contributed substantially to the acquisition of data and provided a critical revision of the article. All authors provided their final approval of the version to publish and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1 Lee LM, Thacker SB. Public health surveillance and knowing about health in the context of growing sources of health data. Am J Prev Med. 2011 Dec;41(6):636–40.
2 Nunes B, Barreto M, Gil AP, Kislaya I, Namorado S, Antunes L, et al. The first Portuguese National Health Examination Survey (2015): design, planning and implementation. J Public Health (Oxf). 2019 Sep;41(3):511–7.
3 Tolonen H, Koponen P, Aromaa A, Conti S, Graff-Iversen S, Grøtvedt L, et al., editors. Review of health examination surveys in Europe. Helsinki, Finland: National Public Health Institute; 2008.
4 Üstun TB, Chatterji S, Mechbal A, Murray CJ. Quality assurance in surveys: standards, guidelines and procedures. United Nations. Department of Economic and Social Affairs. Statistics Division. Household sample surveys in developing and transition countries. New York (NY): United Nations; 2005. pp. 199–230.
5 Kreuter F, editor. Improving surveys with paradata: analytic uses of process information. Hoboken (NJ): John Wiley & Sons; 2013.
6 Groves RM, Lyberg L. Total survey error: past, present, and future. Public Opin Q. 2010;74(5):849–79.
7 Biemer PP. Total survey error: design, implementation, and evaluation. Public Opin Q. 2010;74(5):817–48.
8 Groves RM. Survey errors and survey costs. Wiley Series in Probability and Statistics. Hoboken, NJ: John Wiley & Sons; 1989.
9 Biemer PP, Lyberg LE. Introduction to survey quality. Hoboken (NJ): John Wiley & Sons; 2003.
10 Yan T, Olson K. Analyzing paradata to investigate measurement error. In: Kreuter F, editor. Improving surveys with paradata: analytic uses of process information. Hoboken (NJ): John Wiley & Sons; 2013. pp. 73–95.
11 Santos AJ, Gil AP, Kislaya I, Antunes L, Barreto M, Namorado S, et al. 1o Inquérito Nacional de Saúde Com Exame Físico: INSEF 2015: relatório metodológico. Lisboa: Instituto Nacional de Saúde Doutor Ricardo Jorge; 2016.
12 Tolonen H, Ahonen S, Jentoft S, Kuulasmaa K, Heldal J; European Health Examination Pilot Project. Differences in participation rates and lessons learned about recruitment of participants – the European Health Examination Survey Pilot Project. Scand J Public Health. 2015 Mar;43(2):212–9.
13 Tolonen H, editor. EHES manual: part B: fieldwork procedures. 2013. Helsinki, Finland: National Institute for Health and Welfare; 2013.
14 Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009 Apr;42(2):377–81.
15 Tolonen H, Koponen P, Aromaa A, Conti S, Graff-Iversen S, Grøtvedt L, et al., editors. Recommendations for the health examination surveys in Europe. Helsinki, Finland: National Public Health Institute; 2008.
16 Tolonen H, Dobson A, Kulathinal S; WHO MONICA Project. Assessing the quality of risk factor survey data: lessons from the WHO MONICA Project. Eur J Cardiovasc Prev Rehabil. 2006 Feb;13(1):104–14.
17 Ministério da Saúde; Instituto Nacional de Saúde Doutor Ricardo Jorge. Inquérito de Saúde com Exame Físico: relatório de estudo realizado em São Brás de Alportel. Lisboa: Departamento de Epidemiologia, INSA; 2012.
18 McCaughey EJ, Vecellio E, Lake R, Li L, Burnett L, Chesher D, et al. Key factors influencing the incidence of hemolysis: a critical appraisal of current evidence. Crit Rev Clin Lab Sci. 2017 Jan;54(1):59–72.
19 Dias CM. 25 anos de Inquérito Nacional de Saúde em Portugal. Rev Port Saúde Pública. 2009;Especial 25 Anos:51–60.
20 Mindell JS, Giampaoli S, Goesswald A, Kamtsiuris P, Mann C, Männistö S, et al.; HES Response Rate Group. Sample selection, recruitment and participation rates in health examination surveys in Europe – experience from seven national surveys. BMC Med Res Methodol. 2015 Oct;15(1):78.
21 Blom AG, Julie M. Korbmacher. Measuring interviewer characteristics pertinent to social surveys: a conceptual framework. 2013. Survey methods: insights from the field. Retrieved from: https://surveyinsights.org/?p=817.

5.4. Combining self-reported and objectively measured survey data to improve hypertension prevalence estimates: Portuguese experience

The fourth research paper was published in 2021 in the Archives of Public Health journal. This paper reflects on feasibility of measurement error correction using external validation data. It illustrates application of multiple imputation for correction of measurement error in self-reported hypertension in prevalence and prevalence ratios.

Kislaya et al. Archives of Public Health (2021) 79:45
https://doi.org/10.1186/s13690-021-00562-y

Archives of Public Health

RESEARCH Open Access

Combining self-reported and objectively measured survey data to improve hypertension prevalence estimates: Portuguese experience

Irina Kislaya[1,2,3*], Andreia Leite[2,3], Julian Perelman[2,3], Ausenda Machado[1,2,3], Ana Rita Torres[1], Hanna Tolonen[4] and Baltazar Nunes[1,2,3]

Abstract

Background: Accurate data on hypertension is essential to inform decision-making. Hypertension prevalence may be underestimated by population-based surveys due to misclassification of health status by participants. Therefore, adjustment for misclassification bias is required when relying on self-reports. This study aims to quantify misclassification bias in self-reported hypertension prevalence and prevalence ratios in the Portuguese component of the European Health Interview Survey (INS2014), and illustrate application of multiple imputation (MIME) for bias correction using measured high blood pressure data from the first Portuguese health examination survey (INSEF).

Methods: We assumed that objectively measured hypertension status was missing for INS2014 participants ($n = 13,937$) and imputed it using INSEF ($n = 4910$) as auxiliary data. Self-reported, objectively measured and MIME-corrected hypertension prevalence and prevalence ratios (PR) by sex, age group and education were estimated. Bias in self-reported and MIME-corrected estimates were computed using objectively measured INSEF data as a gold-standard.

Results: Self-reported INS2014 data underestimated hypertension prevalence in all population subgroups, with misclassification bias ranging from 5.2 to 18.6 percentage points (pp). After MIME-correction, prevalence estimates increased and became closer to objectively measured ones, with bias reduction to 0 pp - 5.7 pp. Compared to objectively measured INSEF, self-reported INS2014 data considerably underestimated prevalence ratio by sex (PR = 0.8, 95CI = [0.7, 0.9] vs. PR = 1.2, 95CI = [1.1, 1.4]). MIME successfully corrected direction of association with sex in bivariate (PR = 1.1, 95CI = [1.0, 1.3]) and multivariate analyses (PR = 1.2, 95CI = [1.0, 1.3]). Misclassification bias in hypertension prevalence ratios by education and age group were less pronounced and did not require correction in multivariate analyses.

(Continued on next page)

* Correspondence: irina.kislaya@insa.min-saude.pt
[1]Departament of Epidemiology, National Health Institute Doutor Ricardo Jorge, Lisbon, Portugal
[2]NOVA National School of Public Health, Public Health Research Centre, Universidade NOVA de Lisboa, Lisbon, Portugal
Full list of author information is available at the end of the article

(Continued from previous page)
Conclusions: Our results highlight the importance of misclassification bias analysis in self-reported hypertension. Multiple imputation is a feasible approach to adjust for misclassification bias in prevalence estimates and exposure-outcomes associations in survey data.

Keywords: Hypertension, Self-reports, Misclassification error, Bias correction, Multiple imputation, Survey , MIME

Background

Reliable and precise estimates of hypertension prevalence are essential to inform health policies at the global, regional, national, and local levels [1–3]. High blood pressure prevalence is one of the harmonized European Core Health Indicators continuously monitored at the European Union (EU) and its member states, using self-reported survey data [1, 2].

There is substantial evidence in epidemiological research on limited validity of self-reports to measure hypertension [4–6]. Survey participants could misclassify their health status and report it as healthy due to inaccurate recall or lack of awareness [4–8]. Several studies have shown that, across EU countries, the awareness of hypertension is still far from perfect (ranging between 33.9 and 82.2%) [5–7, 9, 10] and may be differential among population subgroups [5–7, 9], since some are more prone to reporting errors in surveys.

Incorrect reports of hypertension lead to inaccurate survey inference on prevalence and measures of association, so called misclassification bias [4–6, 8], with implications for public health planning, interventions evaluation, and research. To reduce misclassification bias objective measure of blood pressure through health examination surveys (HES) has thus been recommended [6]. However, objective measurements are more expensive and time-consuming, and represent a higher burden for participants [11]. Surveys with objective measurements are usually implemented in a smaller scale with implications for estimates precision, limited level of disaggregation and subsequent limited subpopulation analysis. Even in high-income countries, frequent HES may not be a feasible substitute of large-scale health interview surveys, and as such, decision makers and researchers still rely highly on self-reports. A possible way to address this issue would be to maintain self-report measures from large-scale surveys with high precision but attempt to adjust it for the potential misclassification bias using more accurate HES data.

Several methods, such as regression calibration, maximum likelihood and Bayesian approaches [12–16] and respective software solutions [14, 15] have been put forward to account for misclassification in this context. However, these are complex and might not be intelligible for the average public health researcher. A more feasible alternative is to consider misclassification as a missing data problem and apply multiple imputation techniques for misclassification error correction (MIME) [12, 17–19]. MIME is deemed simple to use and does not require strong programming skills, since multiple imputation routines are included in all standard statistical software [18]. Another advantage of MIME is that it can be applied to correct in a single step bias in both prevalence estimates and exposure-outcome associations, allowing accounting for differential and non-differential misclassification errors either in outcome or exposure, and using internal and external validation data [12, 17–20].

National Health Interview Survey (INS) [21] represents a key tool for monitoring trends in hypertension and other cardiovascular disease risk factors in Portugal, providing evidence for public health planning and National Health Programs [3, 22]. Since 2014, INS is the Portuguese component of European Health Interview Survey, used for European Core Health Indicators monitoring in the EU. In 2015, Portugal developed its first HES combining self-reported and objectively measured data on hypertension for the same individuals, thus providing an opportunity to investigate the magnitude and direction of misclassification bias in self-reports and the feasibility of statistical adjustments.

Despite extensive methodological research on methods performance in a variety of settings and recently growing attention to misclassification bias in self-reports in epidemiological literature, this issue is still overlooked in public health practice; few studies attempted to assess results robustness in the presence of misclassification or adjust for it [23, 24].

This study aims to illustrate the application of multiple imputation for misclassification bias correction in self-reported hypertension prevalence and prevalence ratios in the INS2014, using data on objectively measured blood pressure from the first Portuguese HES (INSEF).

Methods

Study settings

We used data from two population-based surveys: with interview (INS2014) and examination (INSEF) conducted in Portugal in 2014–2015. A detailed description of surveys design, sampling, and data collection has been provided elsewhere [25, 26].

Briefly, INSEF is a cross-sectional study conducted in 2015 by the Instituto Nacional de Saúde Doutor Ricardo Jorge, in partnership with the five Regional Health

Administrations, the Regional Secretariats for Health of the Autonomous Regions (Azores and Madeira) and the Norwegian Institute of Public Health. INSEF collected objectively measured and self-reported data on a multistage probabilistic sample ($n = 4911$, response rate of 43.9%) of non-institutionalized Portuguese population aged 25–74 years old (yo), through physical examinations and interviews.

INS2014 was developed by Statistics Portugal and Instituto Nacional de Saúde Doutor Ricardo Jorge as an integrated part of European Health Interview Survey, wave 2. It is a cross-sectional study targeting non-institutionalized resident population aged 15 years or over. Survey sample ($n = 18,204$) was selected using multistage stratified design, participation rate was 80.8% [25]. INS2014 collected self-reported data on sociodemographic characteristics, health status and its determinants. For this study, INS2014 sample was restricted to individuals aged 25–74 yo ($n = 13,937$), i.e., the age group available in both surveys.

Definitions

Self-reported hypertension prevalence was estimated using the following question of INS2014: "*During the past 12 months, have you had any of the following chronic disease or conditions? High blood pressure/ Hypertension (Yes/No). Consider disease/conditions even if the symptoms were not present due to medical treatment"*. Individuals who answered "yes" were considered hypertensive.

For INSEF, self-reported hypertension prevalence was defined using two questions: i) "*Do you have any of the following longstanding diseases or conditions: High blood pressure or hypertension? (Yes/No)" Consider longstanding disease/conditions which have lasted, or are expected to last, for 6 months or more.* , and if yes, ii) "*Were these conditions diagnosed by a medical doctor? (Yes/No)*". Individuals who answered positively to both questions were considered hypertensive.

Prevalence of objectively measured high blood pressure was defined as the proportion of those: i) with systolic blood pressure ≥ 140 mmHg, or ii) with diastolic blood pressure ≥ 90 mmHg, or iii) reporting to take prescribed medication to control blood pressure in the 2 weeks prior to the interview. Blood pressure was measured in a sitting position after 5 min of rest using automated measurement device OMROM M6. Three sequential blood pressure measurements were taken on the right arm with one-minute intervals. The average of the 2nd and 3rd readings for systolic and diastolic blood pressure was considered. Medication intake was assessed by the questions: "*During the past 2 weeks, have you used any medicines that were prescribed for you by a doctor?*" and if yes, "*Were the medicines for hypertension?*"

Both surveys collected data on sex, age group, region of residence, urbanization, education, income, health behaviours and healthcare use, using similar questions (Table S1).

Statistical analysis

Participants' characteristics and self-reported hypertension prevalence rates were compared between the two surveys using chi-square test. Exploratory analysis of missing data was performed, logistic regression was used to investigate whether probability of objectively measured hypertension being missing is related to observed data.

We used logistic regression imputation method for monotone missing data patterns for misclassification correction. We assumed that objectively measured hypertension was missing at random for INS2014 participants and imputed it using INSEF as auxiliary data. Logistic regression model was fitted on the INSEF sample considering objectively measured hypertension as outcome and self-reported hypertension and set of other covariates as independent variables. Among covariates available in both surveys only statistically significant were included in the final imputation model. Model performance was assessed using area under receiver operating curve (AUC) and Archer-Lemeshow goodness-of-fit test [27]. Fitted model was used to impute "objectively measured" hypertension values in the INS2014 sample. Imputation was based on a set of 100 imputation iterations.

We estimated objectively measured, self-reported and MIME-corrected prevalence of hypertension and respective 95% confidence intervals (95CI) for overall sample and stratified by sex, age group and educational level. Poisson regression models were fitted to estimate prevalence ratios (PR) of self-reported, objectively measured and MIME-corrected hypertension according to sex, age group and educational level. Poisson regression was chosen since it allows to estimate prevalence ratio directly and is recommended as alternative to logistic regression when outcome is not rare [28, 29].

Self-reported and MIME-corrected INS2014 prevalence and prevalence ratios estimates were compared in terms of bias, standard error (SE) and mean squared error (MSE). Bias in self-reported and MIME-corrected estimates were computed using objectively measured INSEF data as gold standard. MSE was estimated as a sum of the variance and the bias squared.

Stata 15.1 was used for data analyses [30]. All analysis presented in the manuscript, including multiple imputation, were performed using sampling weights to account for the complex sample design of INS2014 and INSEF samples. Significance level of 5% was considered.

Results

Participants characteristics and surveys comparability

INS2014 and INSEF respondents were similar regarding sociodemographic characteristics (Table S2), which is consistent with the samples being representative of the Portuguese population. Differences between surveys were observed for three of 11 variables. Notably, we observed a higher proportion of INSEF participants reporting to have their blood pressure measured by a health professional in the last 12 months (82.2% vs. 78.1%), and to consume alcohol in the last 12 months (80.1% vs. 73.1%). In contrast, a lower proportion of INSEF participants reported to have consulted a general practitioner in the last 12 months (65.2% vs.74.9%).

Imputation model

Of potential 11 covariates: self-reported hypertension, sex, age group, region of residence, level of education and practice of physical activity (Table S3) logistic regression model used for MIME included six that were statistically significant. Model showed good fit (Archer-Lemeshow goodness-of-fit test p-value = 0.167) and excellent classification accuracy (AUC = 0.92).

Hypertension prevalence

Self-reported prevalence estimates were similar between both surveys, except for the 55–64 age group, where higher estimates (p-value = 0.0288) were obtained for the INSEF sample (Table 1). In both surveys, self-reported prevalence was lower compared to objectively measured; the extent of underestimation varied by population subgroup.

Following MIME correction, prevalence estimates increased substantially and approximated to their objectively measured INSEF counterparts. Overall, MIME reduced misclassification bias from 11.5 pp. to 0.7 pp. (Table 2). Bias reduction was observed in all studied population subgroups, yet, for 55–64 yo the difference between INS2014 and INSEF remained considerable after correction (5.7 pp).

Regarding prevalence estimates precision (Table S4), when comparing SE for MIME-corrected estimates obtained with larger INS2014 sample and objectively measured ones from the smaller INSEF sample, we observed marginal improvements for 6 population subgroups, while for other 4 corrected estimates were less precise. Gains in terms of MSE after MIME correction were observed in all studied subgroups (Table 2).

Prevalence ratios

Prevalence ratios of hypertension according to sex, age group and educational level estimated by Poisson regression are summarized in Table 3.

Table 1 Prevalence of objectively measured, self-reported and MIME-corrected hypertension stratified by sex, age group and educational level in INSEF and INS2014

	INSEF Self-reported		INS2014 Self-reported		INSEF Objectively measured		INS2014 MIME-corrected	
	p	95CI	p	95CI	p	95CI	p	95CI
Overall	25.7	[23.9, 27.5]	24.5	[23.4, 25.7]	36.0	[34.3, 37.7]	35.3	[33.3, 37.3]
Sex								
Female	26.1	[24.0, 28.3]	26.8	[25.2, 28.5]	32.7	[30.1, 35.5]	33.2	[30.8, 35.6]
Male	25.1	[22.1, 28.4]	22.0	[20.5, 23.7]	39.6	[36.5, 42.8]	37.6	[34.4, 40.8]
Age group								
25–44 years[a]	6.0	[4.6, 7.8]	7.1	[6.0, 8.2]	12.1	[9.9, 14.7]	13.1	[10.6, 15.6]
45–54 years	19.8	[16.7, 23.3]	22.5	[20.2, 25.0]	35.8	[31.3, 40.6]	37.4	[33.1, 41.7]
55–64 years	47.1	[41.4, 53.0]	39.8	[37.0, 42.6]	58.4	[51.4, 65.0]	52.7	[48.1, 57.2]
65–74 years	58.9	[53.0, 64.6]	54.3	[51.2, 57.3]	71.3	[65.7, 76.4]	68.3	[63.9, 72.6]
Educational level								
ISCED[b] 2011 levels 0–1	41.8	[38.2, 45.4]	39.3	[37.3, 41.2]	55.4	[51.9, 58.8]	53.4	[50.0, 56.8]
ISCED 2011 level 2	18.7	[15.4, 22.6]	18.3	[16.0, 20.9]	29.4	[25.0, 34.2]	29.5	[25.1, 33.9]
ISCED 2011 levels 3–4	15.2	[12.2, 18.8]	12.9	[10.9, 15.1]	24.0	[20.6, 27.8]	21.6	[17.4, 25.9]
ISCED 2011 levels 5–8	10.6	[7.8, 14.11]	10.2	[8.4, 12.3]	15.5	[12.6, 18.9]	15.5	[12.3, 18.6]

INSEF – Portuguese Health Examination Survey, INS2014 – National Health Interview Survey 2014, MIME – Misclassification error correction
[a]the age groups 25–34 years and 34–44 years were aggregated in further analyses given a low number of participants reported hypertension
[b]Educational level was grouped according the 2011 International Standard Classification of Education (ISCED) into four categories

Table 2 Misclassification bias and mean squared error (MSE) for self-reported and MIME-corrected hypertension prevalence

	INS2014 Bias, self-reported (pp)	INS2014 Bias, MIME-corrected (pp)	INS2014 MSE, self-reported	INS2014 MSE, MIME-corrected
Overall	11.5	0.7	0.0132	0.0002
Sex				
Female	5.9	−0.5	0.0036	0.0002
Male	17.6	2.0	0.0309	0.0007
Age group				
25–44 years	5.0	−1.0	0.0026	0.0003
45–54 years	13.3	− 1.5	0.0179	0.0007
55–64 years	18.6	5.7	0.0347	0.0038
65–74 years	17.0	3.1	0.0291	0.0014
Educational level				
ISCED 2011 levels 0–1	16.1	2.0	0.0261	0.0007
ISCED 2011 level 2	11.1	−0.1	0.0124	0.0006
ISCED 2011 levels 3–4	11.1	2.4	0.0125	0.0011
ISCED 2011 levels 5–8	5.2	0.0	0.0028	0.0003

MSE = variance + bias^2
Bias is defined as the absolute difference between the INS2014 and INSEF objectively measured prevalence estimates in percentage points (pp)
INSEF – Portuguese Health Examination Survey, INS2014 – National Health Interview Survey 2014, MIME – Misclassification error correction

Comparing men and women based on self-reported INS2014 data, the prevalence ratio was 0.8 [95CI: 0.7, 0.9], indicating lower hypertension prevalence among men. Objectively measured INSEF data pointed out in opposite direction, indicating 1.2-fold increase in prevalence among men (PR = 1.2 [95CI: 1.1, 1.4]), when compared to women. MIME correction produced statistically significant PR estimate of 1.1 [95CI: 1.0, 1.3], closer to the objectively measured.

INS2014 data indicated a 7.7-fold increase in hypertension prevalence [95CI: 6.6, 9.1] in the 65–74 yo group, compared to the reference group (25–44 yo), while a smaller effect of age was estimated with objectively measured INSEF data (PR = 5.9 [95CI: 4.9, 7.2]). Accounting for outcome misclassification through MIME correction produced a PR estimate of 5.2 [95CI: 4.3, 6.4].

Table 3 Prevalence ratios (PR) and corresponding 95% confidence intervals (95CI) of self-reported, objectively measured and MIME-corrected hypertension by sex, age group and educational level in INS2014 and INSEF

	INSEF, self-reported PR[95CI]	INS2014, self-reported PR[95CI]	INSEF, objectively measured PR[95CI]	INS2014, MIME-corrected PR[95CI]
Sex				
Female	ref	ref	ref	ref
Male	1.0 [0.8, 1.1]	**0.8** [0.7, 0.9]	**1.2** [1.1, 1.4]	**1.1** [1.0, 1.3]
Age group				
25–44 years	ref	ref	ref	ref
45–54 years	**3.3** [2.4, 4.6]	**3.2** [2.7, 3.9]	**3.0** [2.4, 3.7]	**2.9** [2.3, 3.6]
55–64 years	**7.9** [5.7, 10.9]	**5.7** [4.8, 6.7]	**4.8** [3.8, 6.1]	**4.0** [3.3, 5.0]
65–74 years	**9.8** [6.9, 14.0]	**7.7** [6.6, 9.1]	**5.9** [4.9, 7.2]	**5.2** [4.3, 6.4]
Educational level				
ISCED 2011 levels 0–1	**4.0** [3.0, 5.3]	**3.8** [3.2, 4.6]	**3.6** [3.0, 4.3]	**3.5** [2.8, 4.3]
ISCED 2011 level 2	**1.8** [1.3, 2.4]	**1.8** [1.4, 2.2]	**1.9** [1.5, 2.5]	**1.9** [1.5, 2.5]
ISCED 2011 levels 3–4	**1.4** [1.0, 2.0]	**1.3** [1.0, 1.6]	**1.6** [1.3, 1.9]	**1.4** [1.1, 1.8]
ISCED 2011 levels 5–8	ref	ref	ref	ref

INSEF – Portuguese Health Examination Survey, INS2014 – National Health Interview Survey 2014, MIME – Misclassification error correction

INS2014 estimates of prevalence ratios according to educational level were similar to the INSEF, indicating no need for correction. Prevalence ratios from MIME correction remained close to the original ones (Table 3).

After adjustment for confounding, we observed misclassification bias required correction in self-reported prevalence ratio of hypertension by sex but not by age group or educational level. MIME-corrected PR of 1.2 [95CI: 1.0, 1.3] resulting from multivariate Poisson regression (Table S5) was similar to objectively measured in INSEF.

Discussion

In this study, we successfully applied MIME correction to adjust self-reported hypertension prevalence and prevalence ratios estimates from INS2014, a large-scale population-based study using external data on objectively measured hypertension from the first Portuguese HES as reference.

Our results showed that self-reported data on hypertension, European Core Health Indicator collected by the Portuguese component of the European Health Interview Survey wave 2, is subject to differential misclassification, and, if ignored, it leads to inaccurate inference and misleading scientific conclusions. Notably, for Portuguese aged 25–74 years old, the underestimated prevalence and severity of bias varied among subgroups. Misclassification bias in prevalence estimates were larger among men, older age groups, and less educated people. Such differences might be explained by recall bias, confusion between "controlled" disease and "cure" and differential health-seeking/contact behaviors (e.g. women have more contact with healthcare services during key life events such as pregnancy and thus might be more aware of high blood pressure) [6–8]. While an understanding of the reasons behind these differences is beyond the scope of our work, a more throughout investigation can shed further light regarding which groups are at higher risk of not being correctly diagnosed.

As expected, MIME approach markedly reduced misclassification bias in overall and strata-specific prevalence estimates. INS2014 MIME-corrected estimates were similar to the objectively measured INSEF prevalence for overall sample (35.3% vs. 36.0%) and in all but the 55–64 yo population subgroups. This is likely to be related to the magnitude of bias, with more heavily biased estimates being harder to correct. Furthermore, we identified and accounted for bias in exposure-outcome associations. It has been previously shown that in the presence of differential misclassification of the outcome the measures of association may be biased in any direction (away or towards null) [14]. In our study, compared to objectively measured, INS2014 self-reported data yielded similar prevalence ratio estimates by education level, but overestimated association with age (7.7 vs. 5.9) and considerably underestimated association with sex (0.8 vs. 1.2). In multivariate analysis, only prevalence ratio by sex required correction (0.8 vs. 1.2). MIME approach successfully corrected direction of association with sex in bivariate and multivariate analysis and also corrected magnitude of the associations with age, where bias were less pronounced. It should be noted that smaller bias in multivariate analysis for the self-reported data might be related to distinct directions of bias to the included variables. Overall, these results are in line with previous research, that reported comparable performance of MIME for prevalence estimates and associations correction with both internal and external validation data [19, 20].

INS2014 sample size was approximately 3-fold the INSEF sample, we thus expected MIME-corrected estimates to be more precise than those derived from objectively measured INSEF data. However, we achieved little or no gain in estimates precision. MIME yielded 12–39% smaller SE for prevalence estimates stratified by sex and age group, whereas for prevalence rates by educational level MIME-corrected SE were 3–11% higher compared to INSEF sample. Simulation studies indicate that gains in precision and MSE depend both on the relative sample size of survey and validation datasets and quality of the multiple imputation models [12, 17, 20]. As recommended in literature, we included in the model potential risk factors for hypertension and variables related to misclassification process in the Portuguese context; our model showed excellent discriminating accuracy (AUC = 0.92). However, we were not able to include other, potentially relevant variables (e.g., body mass index), not present in both surveys or not collected with comparable questions and therefore which could not be included in the imputation model. Inclusion of these additional variables could improve MIME performance. In addition, the sample size ratio between two surveys was small, which might also explain the low gain in precision. More pronounced improvements in precision have been reported in the USA, with 17-fold sample size ratio between NHANES and NHIS [20]. Our results were similar to Edwards et al. [17], who used data with 3.4-fold sample size ratio. Although MIME yielded a small gain in precision, we observed considerably lower MSE for corrected estimates than in the original self-reported data due to bias reduction. Even with little or no gain in precision, bias correction in survey estimates may be important for evidence-based decision-making and public health planning. This is particularly relevant when bias is large or changes the direction of the associations, as in our case, where self-reported data completely distorted sex differences in hypertension prevalence, indicating higher amount of disease among women.

Whilst research may be interested solely in prevalence estimates, most frequently the interest is in exposure-outcome associations. MIME has been previously used in different study designs to correct odds ratios, risk ratios, and hazard ratios [15, 17]. In our study, MIME correction was successfully applied to prevalence ratios estimated by Poisson regression. This result corroborates the flexibility of the MIME approach, applicable to several model types. Another advantage of this correction method is that it allows to account for complex sample design [30]. Nowadays, almost all large-scale surveys use complex sample features (stratification, clustering) to reduce data collection cost. Furthermore, multiple imputation might be extended to account simultaneously for misclassification error and missing data [19]. In our study, the proportion of missing data in covariates used for imputation was below 5%, so we excluded item-missing data. Nevertheless, this might be an important aspect to consider in other studies.

Our approach had several limitations. First, we assumed that information on objectively measured hypertension in INS2014 was missing at random (Table S6), and that missingness was not dependent on individuals' hypertension status. While we cannot formally assess this, a possible selection bias related to hypertension status were minimized by survey design and data collection method, given that INS2014 is representative of Portuguese population [25].

Second, we assumed that misclassification error properties and imputation model are transportable between surveys. Transportability is a critical issue and it arises regardless of the method used to correct estimates for measurement error [12, 20, 31]. When validation data does not correctly reflect the relationship between self-reported and the objectively-measured, more severe bias may be introduced by correction [31]. We extensively investigated surveys comparability for a large number of covariates and demonstrated that proportion of different categories of participants were similar between them. Although the reference period for INSEF and INS2014 survey questions for self-reported hypertension was different *("past 12 months"* vs. *"Do you have any long-standing…")* self-reported hypertension prevalence and prevalence ratio estimates were also comparable, which is reassuring.

Third, data collection settings may represent an additional source of bias, affecting the transportability of imputation model. In INSEF, interviews and examinations were conducted by health professionals in primary healthcare facilities, while in INS2014 most participants were interviewed at home (93.1%) and a small proportion participated via self-administrated questionnaire filled in a web application (6.9%) [25]. We were not able to take these differences into account.

Finally, our approach may be applied only if individual-level data on misclassification and good predictors for multiple imputation model are available. If this is not the case and research has no reliable auxiliary data on misclassification error, alternative methods have been proposed [15].

Although external validation data has been successfully used for self-reported estimates correction in different contexts [16, 20, 32], it may be more reliable to use internal validation data for MIME. Therefore, we recommend addressing the issue of self-reported misclassification by conducting a large-scale survey that incorporates collection of objectively measured blood pressure for a random subset of participants. A combined survey will benefit from a large sample size, precision and representativeness of estimates on required levels of disaggregation and, at the same time, provide relevant information on misclassification error associated to self-report, making possible its analysis and correction reducing transportability issues. It also requires less preparation that two separated surveys, thus increasing efficiency. Selection of predictors for the imputation model should be done for each particular case, as misclassification error properties verified in the present study may not hold for forthcoming European Health Interview Survey waves and other surveys in Portugal, other target populations or settings. National and regional HES carried out in several EU member states in the last decades have demonstrated that direction and severity of misclassification bias in self-reported hypertension vary across place [5, 6, 9] and time [7] and thus might seriously compromise comparisons of European Core Health Indicator between EU member states, regardless of efforts to produce comparable health information. Moreover, reliance on hypertension self-report might affect monitoring of time trends and evaluation of health interventions. Portugal and many other EU countries consider cardiovascular disease as priority for action [22]. When health programs targeting high blood pressure and cardiovascular disease risk factors are in place it is reasonable to expect that awareness of risk factors, in particular hypertension, increases over time with programs implementation, leading to changing misreporting patterns. Therefore, collection of objectively measured blood pressure using standardized measurement protocols should be continued to reassess misclassification bias in the European Core Health Indicator and adjust for it in the further research. The method now proposed can be easily extended to other health indicators subject to misclassification for which a gold-standard measures are feasible to be collected by HES, such us obesity, elevated cholesterol levels, smoking, anemia, among others.

Conclusion

In conclusion, our results support previous research questioning accuracy of self-reported hypertension to estimate hypertension prevalence and exposure-outcome associations in general population and highlight the importance of bias analysis when using self-reported data on hypertension. MIME approach may be useful to assess the robustness of the research conclusions and correct for bias in this instance.

Abbreviations

AUC: Area under receiver operating characteristic curve; EU: European Union; HES: Health examination survey; GP: General practitioner; 95CI: 95% confidence interval; INS: National Health Interview Survey; INSEF: Portuguese Health Examination Survey; ISCED 2011: 2011 International Standard Classification of Education; MIME: Multiple imputation for misclassification error; MSE: Mean squared error; NHANES: The USA National Health and Nutrition Examination Survey; NHIS: The USA National Health Interview Survey; pp: Percentage points; PR: Prevalence ratio; SE: Standard error

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13690-021-00562-y.

Additional file 1: Table S1. Common variables in INS2014 and INSEF used in the study. **Table S2**. Sociodemographic characteristics of INSEF and INS2014 participants aged 25–74 years old. **Table S3**. AUC for imputation logistic regression models. **Table S4**. Standard error (SE) for INS2014 self-reported, MIME-corrected and INSEF objectively measured hypertension prevalence estimates. **Table S5**. Self-reported, objectively measured and MIME-corrected adjusted prevalence ratios of hypertension according to sex, age group and educational level. **Table S6**. Coefficients of logistic regression model for the probability of examination-based hypertension being missing.

Acknowledgements

The authors are grateful to all the professionals that were involved in the INSEF and INS2014 fieldwork and to all the INSEF and the INS2014 participants.

Authors' contributions

IK, AM, AL and BN conceptualized the study. IK performed data analysis and prepared first draft of the manuscript. AM, HT, AL, ART, JP, BN contributed to the interpretation of the results and the critical revision of the manuscript. All the authors read and approved the final version of the manuscript.

Funding

No specific funding was received for this study. The Portuguese National Health Examination Survey 2013–2017 (INSEF) was developed as part of the Pre-defined project of the Public Health Initiatives Program, "Improvement of epidemiological health information to support public health decision and management in Portugal. Towards reduced inequalities, improved health, and bilateral cooperation", that benefits from a 1.500.000€ Grant from Iceland, Liechtenstein and Norway, through the EEA Grants.

Availability of data and materials

Access to the micro data for the INS2014, Portuguese component of the European Health Interview Survey 2, can be requested from Eurostat via a research contract. INSEF data may be provided upon reasonable request and with permission of Ethical Committee of the Instituto Nacional de Saúde Dr. Ricardo Jorge.

Declarations

Ethics approval and consent to participate

INSEF was approved by the National Commission for Data Protection, by the Ethical Committees of the Instituto Nacional de Saúde Dr. Ricardo Jorge and ethical committees of all project partners. All participants provided written inform consent. The INS2014 was developed in accordance with the principles of statistical confidentiality under the Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics.

Consent for publication

Not applicable, there are no details on individual participants within the manuscript.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Departament of Epidemiology, National Health Institute Doutor Ricardo Jorge, Lisbon, Portugal. [2]NOVA National School of Public Health, Public Health Research Centre, Universidade NOVA de Lisboa, Lisbon, Portugal. [3]Comprehensive Health Research Center (CHRC), Universidade NOVA de Lisboa, Lisbon, Portugal. [4]Department of Public Health and Welfare, Finnish Institute for Health and Welfare (THL), Helsinki, Finland.

Received: 15 October 2020 Accepted: 15 March 2021
Published online: 08 April 2021

References

1. Verschuuren M, Gissler M, Kilpeläinen K, Tuomi-Nikula A, Sihvonen A-P, Thelen J, Gaidelyte R, Ghirini S, Kirsch N, Prochorskas R, Scafato E, Kramers P, Aromaa A. Public health indicators for the EU: the joint action for ECHIM (European Community Health Indicators & Monitoring). Arch Public Health. 2013;71(1):12. https://doi.org/10.1186/0778-7367-71-12.
2. Bogaert P, Van Oyen H. An integrated and sustainable EU health information system: national public health institutes' needs and possible benefits. Arch Public Heal. 2017;75(1):3. https://doi.org/10.1186/s13690-016-0171-7.
3. Direção Geral da Saúde. Plano Nacional De Saúde. Revisão e Extensão a 2020. Lisboa; 2015.
4. Paalanen L, Koponen P, Laatikainen T, Tolonen H. Public health monitoring of hypertension, diabetes and elevated cholesterol: comparison of different data sources. Eur J Pub Health. 2018.
5. Gonçalves VSS, Andrade KRC, Carvalho KMB, Silva MT, Pereira MG, Galvao TF. Accuracy of self-reported hypertension: a systematic review and meta-analysis. J Hypertens. 2018;36(5):970–8 [cited 2019 Jun 11]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29232280.
6. Tolonen H, Koponen P, Mindell JS, Männistö S, Giampaoli S, Dias CM, et al. Under-estimation of obesity, hypertension and high cholesterol by self-reported data: comparison of self-reported information and objective measures from health examination surveys. Eur J Pub Health. 2014;24(6):941–8. https://doi.org/10.1093/eurpub/cku074.
7. Mills KT, Bundy JD, Kelly TN, Reed JE, Kearney PM, Reynolds K, et al. Global disparities of hypertension prevalence and control. Circulation. 2016;134(6):441–50 [cited 2020 Sep 16]. Available from: https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.115.018912.
8. Kislaya I, Tolonen H, Rodrigues AP, Barreto M, Gil AP, Gaio V, Namorado S, Santos AJ, Dias CM, Nunes B. Differential self-report error by socioeconomic status in hypertension and hypercholesterolemia: INSEF 2015 study. Eur J Pub Health. 2019;29(2):273–8. https://doi.org/10.1093/eurpub/cky228.
9. Polonia J, Martins L, Pinto F, Nazare J. Prevalence, awareness, treatment and control of hypertension and salt intake in Portugal. J Hypertens. 2014;32(6):1211–21. https://doi.org/10.1097/HJH.0000000000000162.
10. Rodrigues AP, Gaio V, Kislaya I, Graff-Iversen S, Cordeiro E, Silva AC, Namorado S, Barreto M, Gil AP, Antunes L, Santos A, Miguel JP, Nunes B, Dias CM, INSEF Research group. Sociodemographic disparities in hypertension prevalence: results from the first Portuguese national health examination survey. Rev Port Cardiol. 2019;38(8):547–55. https://doi.org/10.1016/j.repc.2018.10.012.
11. Tolonen H, Koponen P, Aromaa A, Conti S, Liv SG, Mark G, et al. Recommendations for the Health Examination Surveys in Europe. National Public Health Institute, editor: Julkaisija-Utgivare-Publisher; 2008.
12. Ni J, Dasgupta K, Kahn SR, Talbot D, Lefebvre G, Lix LM, Berry G, Burman M, Dimentberg R, Laflamme Y, Cirkovic A, Rahme E. Comparing external and internal validation methods in correcting outcome misclassification bias in

logistic regression: a simulation study and application to the case of postsurgical venous thromboembolism following total hip and knee arthroplasty. Pharmacoepidemiol Drug Saf. 2019;28(2):217–26. https://doi.org/10.1002/pds.4693.
13. Lash TL, Fox MP, Maclehose RF, Maldonado G, Mccandless LC, Greenland S. Good practices for quantitative bias analysis. Int J Epidemiol. 2014;43(6): 1969–85. https://doi.org/10.1093/ije/dyu149.
14. Keogh RH, Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: part 1—basic theory and simple methods of adjustment. Stat Med. 2020;39(16):2197–231 [cited 2020 Aug 14]. Available from: https://pubmed.ncbi.nlm.nih.gov/32246539/.
15. Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, Keogh RH, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: part 2—more complex methods of adjustment and advanced topics. Stat Med. 2020;39(16):2232–63 [cited 2020 Aug 14]. Available from: https://pubmed.ncbi.nlm.nih.gov/32246531/.
16. Mentz G, Schulz AJ, Mukherjee B, Ragunathan TE, Perkins DW, Israel BA. Hypertension: development of a prediction model to adjust self-reported hypertension prevalence at the community level. BMC Health Serv Res. 2012;12(1):312. https://doi.org/10.1186/1472-6963-12-312.
17. Edwards JK, Cole SR, Troester MA, Richardson DB. Accounting for misclassified outcomes in binary regression models using multiple imputation with internal validation data. Am J Epidemiol. 2013;177(9):904–12. https://doi.org/10.1093/aje/kws340.
18. Cole SR, Chu H, Greenland S. Multiple-imputation for measurement-error correction. Int J Epidemiol. 2006;35(4):1074–81 [cited 2019 Jul 19]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16709616.
19. Livingston MD, Cannell B, Muller K, Komro KA. Comparing methods of misclassification correction for studies of adolescent alcohol use. Am J Drug Alcohol Abuse. 2018;44(2):160–6 [cited 2019 Jul 19]. Available from: https://www.tandfonline.com/doi/full/10.1080/00952990.2017.1421212.
20. Schenker N, Raghunathan TE, Bondarenko I. Improving on analyses of self-reported data in a large-scale health survey by using information from an examination-based survey. Stat Med. 2010;29(5):533–45. https://doi.org/10.1002/sim.3809.
21. Dias CM. 25 anos de Inquérito Nacional de Saúde em Portugal. Rev Port Saúde Pública. 2009:51–60.
22. Direção-Geral da Saúde. Programa nacional para as doenças cérebro-cardiovasculares 2017. Lisboa; 2017.
23. Brakenhoff TB, Mitroiu M, Keogh RH, Moons KGM, Groenwold RHH, van Smeden M. Measurement error is often neglected in medical literature: a systematic review. J Clin Epidemiol. 2018;98:89–97 Elsevier USA. [cited 2020 Sep 4]. Available from: https://pubmed.ncbi.nlm.nih.gov/29522827/.
24. Ranker LR, Petersen JM, Fox MP. Awareness of and potential for dependent error in the observational epidemiologic literature: a review. Ann Epidemiol. 2019;36:15–19.e2 Elsevier Inc. [cited 2020 Sep 4]. Available from: https://pubmed.ncbi.nlm.nih.gov/31402082/.
25. Instituto Nacional de Estatística. Inquerito Nacional de Saúde 2014. Lisboa: INE; 2016. p. 310. Available from: www.ine.pt
26. Nunes B, Barreto M, Gil AP, Kislaya I, Namorado S, Antunes L, et al. The first Portuguese National Health Examination Survey (2015): design, planning and implementation. J Public Health (Bangkok). 2018; [cited 2018 Sep 30]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/30239797.
27. Archer KJ, Lemeshow S. Goodness-of-fit test for a logistic regression model fitted using survey sample data. Stata J. 2006;6(1):97–105. https://doi.org/10.1177/1536867X0600600106.
28. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol. 2003;3(1):21. https://doi.org/10.1186/1471-2288-3-21.
29. Tamhane AR, Westfall AO, Burkholder GA, Cutter GR. Prevalence odds ratio versus prevalence ratio: choice comes with consequences. Stat Med. 2016; 35(30):5730–5 [cited 2021 Feb 10]. Available from: /pmc/articles/PMC5135596/.
30. StataCorp. Stata statistical software: release 15. College Station: StataCorp LP; 2017.
31. Lohr SL, Raghunathan TE. Combining survey data with other data sources. Stat Sci. 2017;32(2):293–312. https://doi.org/10.1214/16-STS584.
32. Drieskens S, Demarest S, Bel S, De Ridder K, Tafforeau J. Correction of self-reported BMI based on objective measurements: a Belgian experience. Arch Public Heal. 2018;76(1):16–9.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Material

Table S1. Common variables in INS2014 and INSEF used in the study

Variable	INS2014	INSEF
Sex	Male/Female	Male/Female
Age group	Age of participant in complete years at the moment of interview coded into one of 5 age groups 25-34, 35-44, 45-54, 55-64, 65-74	Age of participant in complete years calculated from birth and interview dates was coded into one of 5 age groups: 25-34, 35-44, 45-54, 55-64, 65-74
Education level	What is the highest level of education that you have completed successfully? (No formal education, Basic 1º cicle ou 2º cicle, Basic 3º cicle, Secondary education, Post-secondary education, Higher education including post-graduate); Other Recoded according ISCED-11	What is the highest level of education that you have completed successfully? (No formal education, Basic 1º cicle, Basic 2º cicle, Basic 3º cicle, Secondary education, Post-secondary education, Higher education, Post-graduate (Mcs/PhD), Other Recoded according ISCED-11
BP measurement	When was the last time you have your blood pressure been measured by health professional? (Within the past 12 months, 1 to less than 3 years, 3 to less than 5 years, More than 5 years, Never)	When was the last time you have your blood pressure been measured by health professional? (Within the past 3 months, 3-5 months ago, 6-11 months ago, 12 or more months ago, Never)
GP consultation	When was the last time you consulted GP or family doctor on your behalf? (Less than 12 months ago/ More than 12 months ago/Never)	When was the last time you consulted GP or family doctor in public primary care? (Consider consultations on your behalf only). (Less than 12 months ago/ More than 12 months ago/Never)
Income	What is the approximate net monthly income (after tax and	Would you tell me which group represents your household's total

	other deductions) of all persons in your household? Consider all regular monetary souses of income (employment, social benefits, investments etc). Recoded in quintiles	net monthly income from all these sources after tax and other deductions? Recoded in quintiles
Region of residence	Variable provided in the sampling frame (North, Centre, Lisbon and Tagus Vale, Alentejo, Algarve, Madeira, Azores)	Variable provided in the sampling frame (North, Centre, Lisbon and Tagus Vale, Alentejo, Algarve, Madeira, Azores)
Urbanization	Variable provided in the sampling frame (Rural, Urban, Semiurban)	Variable provided in the sampling frame (Rural, Urban, Semiurban)
Smoking	Do you smoke? (Yes daily, Yes occasionally, No)	Do you smoke? (Yes daily, Yes occasionally, No)
Alcohol consumption	In the last 12 months how often have you had an alcoholic drink of any kind (beer, wine, spirits, liquor, coctails, ets)? (Daily or almost daily, 5-6 days per week, 3-4 days per week, 1-2 days per week, 2-3 days in a month, once a month, less than once a month, Not in last 12 month, I do no drink alcohol, Never in my whole life)	During the past 12 months, did you drink any alcohol (wine, beer, liquor, brandy, etc? (Yes/No)
Physical activity	In a typical week, on how many days do you carry out physical activity for at least 10 min continuously? (indicate 0 if not practicing)	In a typical week, do you engage in any regular physical activity and if yes, for how many days?

In both surveys level of education was grouped according the 2011 International Standard Classification of Education (ISCED) into four categories: ISCED 0-1 level (No formal education/Basic(1 cycle)/ Basic 2 cycle, ISCED 2 level (Basic 3 cycle), ISCED 3-4 levels (Secondary/Pos-secondary), ISCED 5-8 levels (Higher/Post-graduate) education

Table S2. Sociodemographic characteristics of INSEF and INS2014 participants aged 25-74 years old

	INSEF (n=4911)		INS2014 (n=13937)		p-value
	n	%	n	%	
Sex					1.000
Female	2,646	52.5	7,733	52.5	
Male	2,265	47.5	6,204	47.5	
Age group					1.000
25-34 years	714	18.3	2,557	18.3	
25-44 years	1,849	23.5	3,188	23.5	
45-54 years	1,193	22.4	3,023	22.4	
55-64 years	1,098	19.9	3,031	19.9	
65-74 years	771	15.9	2,852	15.9	
Education					0.2215
ISCED 2011 levels 0-1	2,193	40.3	6,824	42.5	
ISCED 2011 level 2	918	18.9	2,310	17.8	
ISCED 2011 levels 3-4	958	21.4	2,353	19.3	
ISCED 2011 levels 5-8	838	19.4	2,450	20.4	
Income					0.5798
1Q (Low)	1,092	19.8	2,784	18.0	
2Q	929	18.4	2,592	18.3	
3Q	872	20.4	2,751	20.1	
4Q	837	19.8	2,833	21.1	
5Q (High)	914	21.7	2,977	22.5	
Last BP measurement					0.0003
Less than 12 months	3,880	82.2	10,673	78.1	
12 month or more or never	895	17.8	3,225	21.9	
Last general practitioner (GP) consultation					<0.0001
Less than 12 months	3,016	65.2	10,126	74.9	
12 month or more	1,812	34.8	3,805	25.1	
Urbanization					0.9419
Rural	1,397	26.4	4,990	26.2	
Urban	3,514	73.6	8,947	73.8	
Region of residence					0.9998
Norte	777	35.4	2,105	35.4	
Centro	706	16.2	2,374	16.3	
LVT	650	34.8	2,354	34.8	
Alentejo	690	4.6	1,587	4.6	
Algarve	644	4.2	1,963	4.2	

RA Madeira	695	2.5	1,870	2.5	
RA Açores	749	2.3	1,684	2.3	
Practice of physical activity at least once a week					0.9238
Yes	1,674	34.2	4,536	34.0	
No	3,235	65.8	9,361	66.0	
Smoking					
Yes	1,115	22.1	3,152	22.6	
No	3,793	77.9	10,775	77.4	
Alcohol consumption in last 12 months					<0.0001
Yes	3,924	80.1	9,511	73.1	
No	986	19.9	4,398	26.9	
p-value for chi-square test to compare participants distribution in INSEF vs. INS2014					

Table S3. AUC for imputation logistic regression models

Variables in the model	AUC	p-value for goodness of fit test
Self-reported hypertension	0.844	1.00
Self-reported hypertension, sex, age group	0.910	<0.001
Self-reported hypertension, sex, age group, region	0.913	0.6644
Self-reported hypertension, sex, age group, region, urbanization	0.915	0.4258
Self-reported hypertension, sex, age group, region, education	0.917	0.6364
Self-reported hypertension , sex, age group, region, education, income	0.920	<0.001
Self-reported hypertension, sex, age group, region, education, BP measurement	0.917	0.1716
Self-reported hypertension, sex, age group, region, education, GP visit	0.919	<0.001
Self-reported hypertension sex, age group, region, education, smoking	0.918	0.354
Self-reported hypertension, sex, age group, region, education, alcohol	0.918	0.2024
Self-reported hypertension, sex, age group, region, education, physical activity	0.920	0.167
* GP visit, Blood Pressure measurement, urbanization, income, smoking, alcohol consumption were not statistically significant and were not included in the final model		

Table S4. Standard error (SE) for INS2014 self-reported, MIME-corrected and INSEF objectively measured hypertension prevalence estimates

	SE self-reported	SE MIME-corrected	SE INSEF obj. measured

Overall	0.0056	0.0099	0.0099
Sex			
Female	0.0077	0.0134	0.0154
Male	0.0075	0.0154	0.0179
Age group			
25-44 years	0.0056	0.0125	0.0117
45-54 years	0.0114	0.0218	0.0259
55-64 years	0.0134	0.0230	0.0380
65-74 years	0.0144	0.0220	0.0299
Education			
ISCED 2011 levels 0-1	0.0094	0.0156	0.0177
ISCED 2011 level 2	0.0117	0.0236	0.0230
ISCED 2011 levels 3-4	0.0099	0.0205	0.0184
ISCED 2011 levels 5-8	0.0092	0.0164	0.0157

Table S5. Self-reported, objectively measured and MIME-corrected adjusted prevalence ratios of hypertension according to sex, age group and educational level

Adjusted PR	INS2014, self			INSEF, self			INSEF, exam-based			MIME-corrected		
SEX	PR	IC 95		PR	IC 95		PR	IC 95		PR	IC 95	
Sex												
Female	ref			ref			ref			ref		
Male	0.8	0.7	0.9	1.0	0.9	1.1	1.2	1.1	1.4	1.2	1.0	1.3
Age group												
35-44 years	ref			ref			ref			ref		
45-54 years	2.7	2.3	3.3	2.9	2.1	4.9	2.6	2.1	3.3	2.5	2.0	3.1
55-64 years	4.5	3.7	5.3	6.6	4.9	8.8	4.0	3.2	5.1	3.3	2.6	4.1
65-74 years	5.7	4.7	6.7	7.9	5.7	10.8	4.8	3.9	5.8	4.0	3.3	5.0
EDUC												
ISCED 2011 levels 0-1	2.1	1.7	2.5	1.7	1.4	2.2	1.8	1.6	2.2	2.0	1.7	2.7
ISCED 2011 level 2	1.6	1.3	1.9	1.4	1.0	1.8	1.5	1.2	1.9	1.7	1.3	2.1
ISCED 2011 levels 3-4	1.3	1.0	1.6	1.3	0.9	1.8	1.4	1.1	1.8	1.4	1.2	1.9
ISCED 2011 levels 5-8	ref			ref			ref			ref		

Table S6. Coefficients of logistic regression model for the probability of examination-based hypertension being missing

	β	CI 95%		p-value
Sex				
Male	0.004709	-0.07832	0.087734	0.911
Age group				
35-44	-0.02034	-0.15427	0.11358	0.766
45-54	-0.01355	-0.20076	0.133483	0.887
55-64	-0.02651	-0.20736	0.15434	0.774
65-74	-0.03492	-0.249	0.17915	0.749
Education				
ISCED 2011 level 2	-0.15724	-0.28735	-0.02713	0.018
ISCED 2011 levels 3-4	-0.20422	-0.37182	-0.03663	0.017
ISCED 2011 levels 5-8	-0.05384	-0.29486	0.187177	0.661
Region of residence				
Centro	0.016786	-0.16972	0.203291	0.860
LVT	0.013114	-0.30659	0.332822	0.936
Alentejo	0.026791	-0.15359	0.207172	0.771
Algarve	0.011887	-0.22249	0.246262	0.921
RA Madeira	-0.01126	-0.25872	0.236193	0.929
RA Açores	3.33E-05	-0.43758	0.43765	1.000
Self-reported hypertension				
Yes	-0.09367	-0.24033	0.052987	0.210
Urbanization				
Rural	0.023895	-0.19796	0.245746	0.833

6 Discussion

The decision making process in health policies and programs are expected to be evidence and information-based. As the number and diversity of information sources available for population health monitoring increase, so should the knowledge and comprehension of their respective strengths and limitations. The purpose of this thesis was to contribute to the improvement of health information available in Portugal for monitoring of cardiovascular disease risk factors focusing on measurement error in self-reported data on hypertension and hypercholesterolemia.

This study intended to cover existing research gaps on the topic addressing measurement error dimensions through five research objectives (Figure 4):

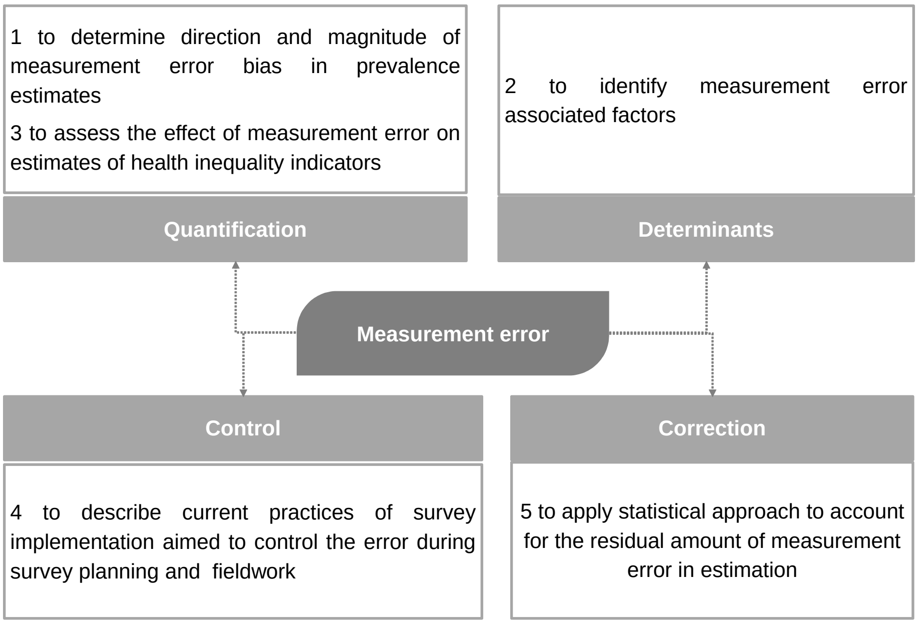


FIGURE 4 MEASUREMENT ERROR DIMENSIONS ADDRESSED THROUGHOUT THE RESEARCH PROJECT

This chapter summarize the main findings, strengths and limitations of the developed research as well as its implications for both further research and practice.

6.1. Main findings

6.1.1. Measurement error quantification

Prior to this study, little was known about measurement error in self-reported data collected in Portugal on hypertension and hypercholesterolemia.

Through this study, I assessed the validity of self-reported data on hypertension and hypercholesterolemia using a large probabilistic sample of the Portuguese population aged 25-74 years old, quantified measurement error bias in prevalence estimates and examined how the measurement error in binary outcome affects outcome-exposure associations in a cross-sectional study.

Considering objectively measured blood pressure and total cholesterol concentrations adjusted for medication intake as a reference, the study showed that self-reported data on hypertension and hypercholesterolemia, collected by a survey for health monitoring in Portugal, is subject to appreciable measurement error, despite all the efforts of careful survey planning and data quality control. Another important finding of this study was that measurement errors in self-reported hypertension and hypercholesterolemia were differential, varied among population subgroups.

The results revealed suboptimal sensitivity of self-reports for both risk factors, suggesting that the reporting error occurred more frequently among individuals who were considered to have hypertension and/or hypercholesterolemia according to the objective measurements. By the other hand, the specificity of self-report in a population under study was high for both risk factors, suggesting that those who do not have hypertension and/or hypercholesterolemia according to the objective measurements rarely misclassified themselves. The distinction between two types of measurement error (over- and underreporting), should be mentioned as one of the added values of this study, as it was novel at national level. Although several studies were developed in Portugal using a combination of questionnaire data and objective measurements of hypertension and hypercholesterolemia (93,112,113,157–159), to our best knowledge, none of those performed detailed investigation of quality of self-reported data from a measurement error perspective, distinguished two error types.

The observed patterns of misreporting led to underestimation of the prevalence of both CVD risk factors, however the degree of underestimation varied, being higher for hypercholesterolemia. Namely, self-reported data underestimated hypertension prevalence by 10.3 percentage points, while for hypercholesterolemia the degree of underestimation was as high as 38.4 percentage points. Comparison of measurement error bias for two CVD risk factors measured on the same sample of participants was also novel at the national level.

This work also illustrated higher agreement between self-reported and objectively measured data on hypercholesterolemia when the higher cut-off value was used for objectively measured indicator definition. Change of operational definitions of indicators leads to a small improvements of estimates.

To assess the effect of measurement error in a binary outcome on the outcome-exposure associations the relative index of inequality (RII) and the slope index of inequality (SII) were estimated. As outlined previously in chapter 4, RII and SII are regression-based indicators used to measure the magnitude of health inequalities and can be interpreted as a prevalence ratio and a prevalence difference between two extremes of the socio-economic distribution, respectively. Moreover, in this thesis, I established a methodological approach that could be used to formally compare health inequality indicators derived using different measurement strategies for health outcomes. Application of GEE models to quantify bias in the RII and SII based on self-reported data was novel at national as well as at European level, as previous research on health inequalities in European countries was limited to qualitative assessment of measurement error bias (160).

Given that self-reported data on outcome is more specific than sensitive, in a presence of a non-differential measurement error, it is expected to observe an attenuation of outcome-exposure associations towards null, however, the direction of bias needs to be carefully investigated case by case when the error is differential(37,41).

Considering educational level as exposure, this study corroborated the hypothesis that in a presence of differential measurement error in a binary outcome, estimates of prevalence ratios and prevalence differences can be biased in any direction, either away or towards null. This finding is particularly relevant and should be considered in the interpretation of data on group comparisons during public health planning and programs evaluation as well as in research on health inequalities in hypertension and hypercholesterolemia when these conditions are measured by self-reports.

6.1.2. Determinants of measurement error

This study developed in Portugal identified modifiable and non-modifiable factors associated with the measurement error in self-reports, focusing on its underreporting component. In specific, for hypertension the observed underreporting error varied by sex, age group, income and education level, while for hypercholesterolemia errors were associated with participants´ age, income and education. In addition, the study's results enhanced the importance of regular contact with primary health care for correct reporting of both hypertension and hypercholesterolemia in surveys. It should be mentioned that, the effect of recent health care use observed in this study was more pronounced for hypertension than for hypercholesterolemia. The results also showed that reporting error occurs mainly to underdiagnoses. These findings may be useful (i) to develop intervention strategies aimed on the improvement of quality of self-reported data on hypertension and hypercholesterolemia and (ii) to identify the priority groups to be targeted by such interventions.

6.1.3. Measurement error control

How to control measurement error? This question has been extensively by survey researchers and epidemiologists. It is well known that whenever possible errors should be limited by design(26). To the successful survey implementation, it is also crucial to perform standardization of all data collection procedures, to establish tangible quality targets through a set of key performance indicators and to closely monitor them during the fieldwork to minimize data collection errors(71). However, there is no standard, “one fits all” solution for measurement error control in survey data.

In this study, I described the main strategies used in the large-scale survey to verify the adherence to the measurement protocols for both self-reported data and objectively measured biomarkers, as well as interventions implemented by fieldwork supervisors to improve the quality of collected data.

The results suggest that interviewer observation and re-interviewing, the principal quality control strategies for self-reported data collection, rely considerably more on subjective judgement of fieldwork supervisors, compared to statistical quality control measures implemented on a large scale for biomarkers. This study emphasized the importance of continuous monitoring of established key performance indicators to ensure data quality throughout the survey. It also showed that continuous monitoring implies additional data gathering effort, however, collected paradata offers the opportunity to make informed real-time decisions on data quality improvements.

6.3.4. Measurement error correction

As the measurement error may persist despite the implemented control strategies, the important question on feasibility of its correction in estimation has arisen. Several methodologies has been proposed to adjust the estimates for measurement error (12,34,161). Although there has been an extensive theoretical developments, and growing attention to this topic in recent research, the public health researchers rarely quantify bias and perform adjustments when using self-reported data.

In this study, I illustrated how to use multiple imputation to assess the robustness of the research conclusions and correct for bias while estimating. Using external data on objectively measured hypertension, I was able to successfully account for bias in prevalence and prevalence ratios estimated in INS 2014 a large-scale self-reported survey used in Portugal as the main data source for ECHI and to monitor progress towards the goals established by the National Health Plan 2020 and Priority Health Programs.

In previous research, multiple imputation was used to adjust for bias coefficients in logistic, binomial and Cox regression models. In this study, the method was applied to correct measurement error bias in prevalence ratios estimated by Poisson regression model. In addition, strengths and limitations of the multiple imputation as bias correction method were investigated. The results suggest that its application is limited to situations when individual-level validation data is available, as well as variables associated with measurement error to make possible adjustment of imputation model.

6.2 Strengths

One of the primary strengths of this work is that it relied on large scale-population-based surveys data. Both INSEF and INS 2014 surveys used in this thesis were developed on multi-stage probabilistic samples of the Portuguese population. Surveys had a good geographical coverage, including both Portugal mainland and autonomous regions and are representative of the target population. In addition, the INS 2014 is an official statistics survey, that was developed under Eurostat guidelines and integrated in a EU statistical system. Moreover, both surveys had large sample size that had its reflection on estimates precision. The considerable sample size of the INSEF survey enabled the examination of measurement error in self-reported hypertension and hypercholesterolemia in several population subgroups, in a level of detail not previously reported for Portuguese population (10).

Other strengths of this work include reliance on international guidelines for definition of hypertension and hypercholesterolemia health indicators, use of standardized measurement protocols promoted by European Centre for Health Examination Surveys, and comply with high quality standards of HES procedures.

6.3. Limitations

The specific limitations of each particular study were outlined in respective chapters. Here, I attempt to focus on general limitations of this work as a whole, related to the data sources and applied methodology that may affect generalizability of the findings.

The first set of limitations to be mentioned is related to the data sources used in this thesis, as it was based on secondary analysis of survey data. Let start with the target population, which includes Portuguese residents aged 25-74 years old, living in the community. Institutionalized individuals, namely those living in nursing homes, prisons and other collective dwellings, were excluded from this study. Although the exclusion of institutionalized is a common feature of official statistics health surveys, it should be kept in mind then generalizing the study results. Additionally, this study focused on population aged 25-74 years, so its findings cannot be generalized to other age groups. Further research is required to investigate measurement error bias in self-reported hypertension and hypercholesterolemia among Portuguese residents younger than 25 or older than 74 years old.

Another limitation of this work, imposed by the data sources, affected investigation of factors associated with measurement error. Literature suggests that measurement error in surveys is a result of interaction between the data collection mode the interviewers' and participants´ characteristics (31,62,162). Considering potential determinants of measurement error, I focused exclusively on participants´ characteristics. The effect of data collection mode on quality of self-reports and the interviewers role remained out of scope of this research. Data on interviewers´ characteristics were not collected originally by the INSEF survey, the attempt to obtain such data after the end of INSEF fieldwork was also not successful. Given that INSEF survey used a single data collection mode (computer-assisted face-to-face interviews) to obtain self-reported data on hypertension and hypercholesterolemia from all survey participants, it was also not possible to make a judgement on the effect of data collection mode on measurement error. These should be kept in mind then generalizing the study conclusions. It also should be acknowledged, that some participant's characteristic identified in literature as potentially associated with reporting error in hypertension and hypercholesterolemia in other countries, such as family history of cardiovascular disease, were not considered in this work due to lack of data.

In addition, among data-related limitations, it should be mentioned that although this study established that overreporting errors were differential among population subgroups, due to small frequency of this type of measurement error in the INSEF sample it was not possible to perform multivariate analysis in order to identified associated factors.

Although it has been already outlined in all previous chapters, it is important to mention here three methodological limitations common in the field of health examination surveys that are related to the measurement strategy in this study. Survey measurements were taken in a single occasion, while for clinical diagnosis blood pressure and cholesterol should be measured in several occasions. Also, total cholesterol concentration was used as a proxy measure to define hypercholesterolemia, the specific fractions (LDL, HDL) were not considered in this study. In addition, for determination of lipid profiles non-fasting samples were used. No fasting requirement was established in INSEF by one hand to maximize participation rates making possible to schedule participation during all working day and by the other hand to reduce the length of fieldwork period. Mentioned factors may lead to overestimation of objectively measured prevalence of hypertension and hypercholesterolemia and overestimation of measurement error bias.

Another methodological limitation, in particular related to quantification of bias in health inequalities indicators and measurement error correction by multiple imputation, is that in outcome-exposure associations the exposure was assumed to be measured without error. Information on basic demographic characteristics such as sex, age, region of residence was already available in the sampling frame and additionally was verified by the interviewers during face-by-face questionnaire administration, so it seemed reasonable to assume that these variables were of good quality. Other variable considered in the study, such as level of education, income, use of health case, were also self-reported. However, there was no attempt made to evaluate the accuracy of their reporting, as no gold-standard data were available to perform measurement error assessment at the time of this project implementation. Further research is required to cover this gap for Portuguese population.

Finally, the assessment of self-reported data validity and correction of bias by multiple imputation in a present study were performed in a single time point. Due to time restrictions of this project, it was not possible to assess whether an established measurement error model for self-reported hypertension and hypercholesterolemia will hold stable over time.

6.3 Implications for research and practice

This work focused on measurement error in self-reported data on hypertension and hypercholesterolemia. Despite several limitations that have been acknowledged in the previous section, the results of this thesis can be useful to improve the information available in Portugal for monitoring of these cardiovascular disease risk factors.

Survey data are extremely valuable for health policy, monitoring of population health and its determinants over time, design and evaluation of public health interventions and cross-country comparisons. Accuracy of self-reported data remains a key element of valid inference. This study showed that for Portuguese population aged 25-74 years old self-reported data from HIS are not optimal to derive estimates of hypertension or hypercholesterolemia prevalence that could have several implications for public health practice and research.

Objectively measured data from HES clearly have an advantage since it captures both diagnosed and yet undiagnosed cases. Actual blood pressure measurements and cholesterol concentrations should be used in clinical studies as well as for public health monitoring in the future. The use of objectively measured data in surveys will improve the accuracy of epidemiological information available for monitoring of cardiovascular disease risk factors. However, the change of preferable data source from HIS to HES for official health statistics is a long journey; it requires time, resources, political consensus. Although several countries have performed national HES, comparable HES data at EU level are still lacking. At national level, several research centers already have experience in biomarkers collection in a large scale to study specific health conditions (92,112,113) as well as experience of national HES and international collaboration in European surveys (21). These installed capacities resulted from national experience and international collaborations may be a starting point to a smooth transition and a wider use of objectively measured data for health monitoring in Portugal.

In mid-term, for population-based surveys, collection of objectively measured data for subsample of respondents may also constitute a feasible solution allowing to assess associated measurement error through sensitivity analysis with a relatively small budget increase. In research, while using self-reported data, the observed socio-economic patterns of measurement error should be considered then interpreting health statistics and research results. Studies that rely on self-reports of hypertension and hypercholesterolemia will benefit from sensitivity analysis, which tests a robustness of findings in a presence of measurement error. It is particularly recommended then study social determinates and health inequalities.

Methods available for sensitivity analysis in a presence of measurement error were mainly developed for data collected under assumption of simple random sampling, there is a gap in methodological and software solutions for a large-scale surveys based on complex samples designs that need to be addressed by further research.

Based on the findings of the current study regarding hypercholesterolemia, it is recommended to reassess the question formulation and its position within survey data collection instrument. Additional qualitative studies are required to improve the quality of self-reported indicators for lipid profile monitoring.

Identification of factors associated with incorrect reporting also constitutes an important step to the improvement of epidemiological information based on self-reports. The current study showed that the inclusion of additional questions on medication intake in survey and change of operational definitions for self-reported hypertension and hypercholesterolemia would lead to a quite small improvement in estimates, given that reporting error occurs mainly due to underdiagnoses and lack of awareness.

Access to healthcare and health literacy were identified in this thesis as key areas for intervention to increase population awareness of hypertension and hypercholesterolemia and improve reporting. Health literacy interventions can be complex and may not have an immediate effect, however they are necessary to change currently existing patterns of healthcare use, highlighting role of preventive care. The design of interventions to improve health literacy and reporting remains out of the scope of this work, so further research is needed to address this gap.

Regarding healthcare access, the increase of population proportion that has personal GP can be beneficial and raise awareness levels, as previous research that showed that those who use the same care provides have a higher probability to undergo preventive screenings and being diagnosed.

Differences in screening opportunities may explain the different degrees of reporting error between hypertension and hypercholesterolemia observed for the Portuguese population. The measurement of blood pressure has lower burden for the participant and can be easily performed in clinical practice by any health professional during any consultation. Measurement of cholesterol is more intrusive in nature and requires greater participant commitment and proactivity to fulfill medical recommendations and perform prescribed testing. In addition, cholesterol testing requires some out of pocket payments. Although small in amount, for some population subgroups these expenditures may constitute an additional barrier to fulfill medical recommendations.

Since the data collection for this thesis several changes occurred in Portugal in the domain of healthcare access, namely, due to great political commitment, the proportion of the Portuguese population with personal GP increased from 86% in 2015 to 92% in 2020 (163). The out of pocket payments associated with primary care were progressively reduced (164) and since 2021 all primary care GP consultations in the public sector of the healthcare system as well as all prescribed complementary diagnostic tests became free of charge for all population subgroups (165). In light of these changes, the accuracy of self-reporting is also expected to change. So, further research is required to understand if the accuracy of self-reported hypertension and hypercholesterolemia changes over time and to investigate the direction and magnitude of such changes. Regular implementation of surveys that combine self-reported and objectively measured data is recommended to monitor population-level change in hypertension and hypercholesterolemia and associated measurement error

7 Conclusions

In Portugal, self-reported data is widely used to inform policy and support National Health Plan and Priority Health Programs. At present self-reports from HIS constitute a basis for hypertension and hypercholesterolemia monitoring.

This thesis aimed to evaluate the accuracy of self-reported data collected by large-scale population-based surveys for monitoring of cardiovascular disease risk factors, focusing on hypertension and hypercholesterolemia. The findings of this thesis advanced the understanding of whether self-reports can provide valid information on cardiovascular diseases risk factors prevalence and its distribution in the Portuguese population. Through analysis of measurement error, quantification of its impact on estimates and application of correction methodology developed studies demonstrated the potential to improve health information quality in Portugal.

Overall, results indicate that self-reported data on hypertension and hypercholesterolemia is far from perfect and contains a substantial amount of error that is unequally distributed among population subgroups. Observed error patterns lead to underestimation of hypertension and hypercholesterolemia prevalence and distortion of health inequalities indicators. A high amount of error due to underdiagnoses particularly in younger age groups rise thoughtful concerns highlighting need for urgent public health interventions to increase population awareness.

Further opportunities to improve reporting quality in surveys rely on strengthening the healthcare system, improving healthcare access and integration of objectively measured data in health surveys on regular basis.

8 References

1. Brownson RC, Fielding JE, Maylahn CM. Evidence-based Decision Making to Improve Public Health Practice. Front Public Heal Serv Syst Res. 2013;2(2):3–9.

2. Verschuuren M, Gissler M, Kilpeläinen K, Tuomi-Nikula A, Sihvonen A-P, Thelen J, et al. Public health indicators for the EU: the joint action for ECHIM (European Community Health Indicators & Monitoring). Arch Public Health. 2013;71:12.

3. Oyebode O, Mindell JS. A review of the use of health examination data from the Health Survey for England in government policy development and implementation. Arch public Heal Arch belges sante publique. 2014;72(1):24.

4. Bogaert P, Van Oyen H. An integrated and sustainable EU health information system: national public health institutes' needs and possible benefits. Arch Public Heal. 2017;75(1):3.

5. Direção Geral da Saúde. Plano Nacional De Saúde. Revisão e Extensão a 2020. Lisboa; 2015.

6. Dias CM. 25 anos de Inquérito Nacional de Saúde em Portugal. Rev Port Saúde Pública. 2009;51–60.

7. European Parliament, European Commission. Regulation (EU) No 282/2014 on the establishment of a third Programme for the Union's action in the field of health (2014-2020). OJ L [Internet]. 2014;86(1):1–13. Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0282&from=EN

8. Official Journal of the European Union, European Parliament, European Commission. Regulation (EU) No 282/2014 of the European Parliament and of the Council of 11 March 2014 on the establishment of a third Programme for the Union's action in the field of health (2014-2020) [Internet]. Official Journal of the European Union 2014 p. 13. Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0282&from=EN

9. Groves RM, Lyberg L. Total survey error: Past, present, and future. Public Opin Q. 2010 Jan;74(5):849–79.

10. Tolonen H, Koponen P, Mindell JS, Männistö S, Giampaoli S, Dias CM, et al. Under-estimation of obesity, hypertension and high cholesterol by self-reported data: Comparison of self-reported information and objective measures from health examination surveys. Eur J Public Health. 2014 Dec;24(6):941–8.

11. Tsai ACH, Chang TL. Quality issues of self-report of hypertension: Analysis of a population representative sample of older adults in Taiwan. Arch Gerontol Geriatr. 2012;55(2):338–42.

12. Schenker N, Raghunathan TE, Bondarenko I. Improving on analyses of self-reported data in a large-scale health survey by using information from an examination-based survey. Stat Med. 2010;29(5):533–45.

13. Vargas CM, Burt VL, Gillum RF, Pamuk ER. Validity of Self-Reported Hypertension in the National Health and Nutrition Examination Survey III , 1988 – 1991. Prev Med (Baltim). 1997;26:678–85.

14. Taylor A, Grande ED, Gill T, Pickering S, Grant J, Adams R, et al. Comparing self-reported and measured high blood pressure and high cholesterol status using data from a large representative cohort study. Aust N Z J Public Health. 2010;34(4):394–400.

15. Huerta JM, José Tormo M, Egea-Caparrós JM, Ortolá-Devesa JB, Navarro C. Accuracy of Self-Reported Diabetes, Hypertension, and Hyperlipidemia in the Adult Spanish Population. DINO Study Findings. Rev Española Cardiol (English Ed. 2009;62(2):143–52.

16. Tenkorang EY, Sedziafa P, Sano Y, Kuuire V, Banchani E. Validity of Self-Report Data in Hypertension Research: Findings From The Study on Global Ageing and Adult Health. J Clin Hypertens (Greenwich). 2015;17(12):977–84.

17. Bago d'Uva T, O'Donnell O, Van doorslaer E. Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. Int J Epidemiol. 2008;37(6):1375–83.

18. Moody A, Cowley G, Ng Fat L, Mindell JS. Social inequalities in prevalence of diagnosed and undiagnosed diabetes and impaired glucose regulation in participants in the Health Surveys for England series. BMJ Open. 2016 Feb 8;6(2):e010155.

19. WHO Regional Office for Europe. Health 2020: A European Policy Framework and Strategy for the 21st Century. World Health Organization; 2013. 190 p.

20. Kuulasmaa K, Tolonen H, Koponen P, Kilpeläinen K, Avdicová M, Broda G, et al. An overview of the European Health Examination Survey Pilot Joint Action. Arch Public Heal. 2012 Dec 28;70(1):20.

21. Tolonen H, Koponen P, Mindell J, Männistö S, Kuulasmaa K. European Health Examination Survey-towards a sustainable monitoring system. Vol. 24,

European Journal of Public Health. 2014. p. 338–44.

22. Tolonen H, Koponen P, Aromaa A, Conti S, Liv SG, Mark G, et al. Recommendations for the Health Examination Surveys in Europe. National Public Health Institute, editor. Julkaisija-Utgivare-Publisher; 2008.

23. Aromaa A, Koponen P, Tafforeau J, Vermeire C, HIS/HES Core Group. Evaluation of Health Interview Surveys and Health Examination Surveys in the European Union. Eur J Public Health. 2003 Sep;13(3 Suppl):67–72.

24. Nunes B, Barreto M, Gil AP, Kislaya I, Namorado S, Antunes L, et al. The first Portuguese National Health Examination Survey (2015): design, planning and implementation. J Public Health (Bangkok) [Internet]. 2018 Sep 19 [cited 2018 Sep 30]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/30239797

25. Last J, editor. A dictionary of epidemiology. 4th ed. Oxford University Press; 2001.

26. Rothman KJ, Greenland S, Lash TL, editors. Modern Epidemiology. 3rd ed.

27. Gordis L. Epidemiology. 5th ed. Elsevier; 2014.

28. Lash TL, Fox MP, Fink AK. Applying Quantitative Bias Analysis to Epidemiologic Data. Media. Springer Science+Business Media, LLC; 2009.

29. Biemer PP. Total survey error: Design, implementation, and evaluation. Public Opin Q. 2010;74(5):817–48.

30. Biemer PP, De Leeuw ED, Eckman S, Edwards B, Kreuter F, Lyberg, Lars E, editors. Total Survey Error in Practice. John Wiley & Sons, Inc.; 2017.

31. Aday LA, Cornelius LJ. Designing and Conducting Health Surveys: A Comprehensive Guide. 3rd ed. John Wiley & Sons; 2011.

32. Buonaccorsi JP. Measurement error : models, methods, and applications. 437 p.

33. Paalanen L, Koponen P, Laatikainen T, Tolonen H. Public health monitoring of hypertension, diabetes and elevated cholesterol: comparison of different data sources. Eur J Public Health. 2018 Feb 16;

34. Mentz G, Schulz AJ, Mukherjee B, Ragunathan TE, Perkins DW, Israel BA. Hypertension: Development of a prediction model to adjust self-reported hypertension prevalence at the community level. BMC Health Serv Res. 2012;12(1):312.

35. Guerard B, Omachonu V, Harvey RA, Hernandez SR, Sen B. The Influence of

Respondent Characteristics on the Validity of Self-Reported Survey Responses [Internet]. Vol. 51, Health Services Research. Health Research & Educational Trust; 2016 [cited 2021 Sep 17]. p. 937–52. Available from: /pmc/articles/PMC4874833/

36. Van Smeden M, Lash TL, Groenwold RHH. Reflection on modern methods: Five myths about measurement error in epidemiological research. Int J Epidemiol [Internet]. 2020 Feb 1 [cited 2020 Aug 26];49(1):338–47. Available from: https://pubmed.ncbi.nlm.nih.gov/31821469/

37. Keogh RH, Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1—Basic theory and simple methods of adjustment. Stat Med [Internet]. 2020 Jul 20 [cited 2020 Aug 14];39(16):2197–231. Available from: https://pubmed.ncbi.nlm.nih.gov/32246539/

38. Lash TL, Fox MP, Maclehose RF, Maldonado G, Mccandless LC, Greenland S. Good practices for quantitative bias analysis. Int J Epidemiol [Internet]. 2014 Dec 1 [cited 2021 Sep 14];43(6):1969–85. Available from: https://academic.oup.com/ije/article/43/6/1969/705764

39. Greenland S. Basic Methods for Sensitivity Analysis of Biases. Int J Epidemiol [Internet]. 1996 Dec 1 [cited 2021 Sep 17];25(6):1107–16. Available from: https://academic.oup.com/ije/article/25/6/1107/672425

40. Brakenhoff TB, Mitroiu M, Keogh RH, Moons KGM, Groenwold RHH, van Smeden M. Measurement error is often neglected in medical literature: a systematic review [Internet]. Vol. 98, Journal of Clinical Epidemiology. Elsevier USA; 2018 [cited 2020 Sep 4]. p. 89–97. Available from: https://pubmed.ncbi.nlm.nih.gov/29522827/

41. Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, Keogh RH, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 2—More complex methods of adjustment and advanced topics. Stat Med [Internet]. 2020 Jul 20 [cited 2020 Aug 14];39(16):2232–63. Available from: https://pubmed.ncbi.nlm.nih.gov/32246531/

42. Edwards JK, Cole SR, Troester MA, Richardson DB. Accounting for misclassified outcomes in binary regression models using multiple imputation with internal validation data. Am J Epidemiol. 2013;177(9):904–12.

43. Cole SR, Chu H, Greenland S. Multiple-imputation for measurement-error correction. Int J Epidemiol [Internet]. 2006 Aug 1 [cited 2019 Jul 19];35(4):1074–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16709616

44. Ni J, Dasgupta K, Kahn SR, Talbot D, Lefebvre G, Lix LM, et al. Comparing external and internal validation methods in correcting outcome misclassification bias in logistic regression: A simulation study and application to the case of postsurgical venous thromboembolism following total hip and knee arthroplasty. Pharmacoepidemiol Drug Saf. 2019;28(2):217–26.

45. Livingston MD, Cannell B, Muller K, Komro KA. Comparing methods of misclassification correction for studies of adolescent alcohol use. Am J Drug Alcohol Abuse [Internet]. 2018 Mar 4 [cited 2019 Jul 19];44(2):160–6. Available from: https://www.tandfonline.com/doi/full/10.1080/00952990.2017.1421212

46. Gilbert R, Martin RM, Donovan J, Lane JA, Hamdy F, Neal DE, et al. Misclassification of outcome in case-control studies: Methods for sensitivity analysis. Stat Methods Med Res. 2016 Oct 1;25(5):2377–93.

47. Tang L, Lyles RH, King CC, Celentano DD, Lo Y. Binary regression with differentially misclassified response and exposure variables. Stat Med. 2015 Apr 30;34(9):1605–20.

48. Lyles RH, Lin J. Sensitivity analysis for misclassification in logistic regression via likelihood methods and predictive value weighting. Stat Med [Internet]. 2010 Sep 30 [cited 2021 Sep 16];29(22):2297–309. Available from: https://pubmed.ncbi.nlm.nih.gov/20552681/

49. Chen Q, Galfalvy H, Duan N. Effects of disease misclassification on exposure-disease association. Am J Public Health [Internet]. 2013 Apr 10 [cited 2021 Sep 17];103(5). Available from: https://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2012.300995

50. Jurek AM, Greenland S, Maldonado G. How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? Int J Epidemiol [Internet]. 2008 Apr [cited 2021 Sep 16];37(2):382–5. Available from: https://pubmed.ncbi.nlm.nih.gov/18184671/

51. Schouten B, Bais F, Toepoel V. Estimating survey questionnaire profiles for measurement error risk. J Surv Stat Methodol [Internet]. 2018 Sep 1 [cited 2021 Sep 17];6(3):306–34. Available from: https://academic.oup.com/jssam/article/6/3/306/4565751

52. Beatty PC, Collins D, Kaye L, Padilla JL, Willis GB, Wilmot A. Advances in questionnaire design, development, evaluation, and testing [Internet]. Advances in Questionnaire Design, Development, Evaluation and Testing. wiley; 2019 [cited 2021 Sep 16]. 1–785 p. Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/9781119263685

53. Robinson SB, Leonard KF. Designing quality survey questions. SAGE Publications; 2019. 229 p.

54. Tipping S, Hope S, Pickering K, Erens B, Roth MA, Mindell JS. The effect of mode and context on survey results: Analysis of data from the Health Survey for England 2006 and the Boost Survey for London. BMC Med Res Methodol 2010 101 [Internet]. 2010 Sep 27 [cited 2021 Sep 16];10(1):1–8. Available from: https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-10-84

55. Beebe TJ, Ziegenfuss JY, Jenkins SM, Lackore KA, Johnson TP. Survey mode and asking about future intentions did not impact self-reported colorectal cancer screening accuracy. BMC Med Res Methodol [Internet]. 2014 Feb 5 [cited 2021 Sep 16];14(1). Available from: https://pubmed.ncbi.nlm.nih.gov/24499399/

56. Clement SL, Severin-Nielsen MK, Shamshiri-Petersen D. Device effects on survey response quality. A comparison of smartphone , tablet and PC responses on a cross sectional probability sample. Surv Methods Insights from F [Internet]. 2020 [cited 2021 Sep 16];(Special issue: ‘Advancements in Online and Mobile Survey Methods’):186–8. Available from: https://surveyinsights.org/?p=13585

57. Hoebel J, von der Lippe JE, Lange C, Ziese T. Mode differences in a mixed-mode health interview survey among adults. Arch Public Heal [Internet]. 2014 [cited 2021 Sep 16];72(1). Available from: https://pubmed.ncbi.nlm.nih.gov/25810913/

58. Olson K, Smyth JD, Dykema J, Holbrook AL, Kreuter F, West BT. Interviewer Effects from a Total Survey Error Perspective. Interviewer Effects from a Total Survey Error Perspective. Chapman and Hall/CRC; 2020.

59. Davis RE, Couper MP, Janz NK, Caldwell CH, Resnicow K. Interviewer effects in public health surveys. Health Educ Res [Internet]. 2010 Feb [cited 2021 Sep 16];25(1):14–26. Available from: https://pubmed.ncbi.nlm.nih.gov/19762354/

60. Blom AG, Julie M. Korbmacher. Measuring Interviewer Characteristics Pertinent to Social Surveys: A Conceptual Framework. Surv Methods Insights from F.

2013;1–16.

61. West BT, Blom AG. Explaining interviewer effects: A research synthesis. J Surv Stat Methodol [Internet]. 2017 [cited 2021 Sep 16];5(2):175–211. Available from: https://academic.oup.com/jssam/article/5/2/175/2452318

62. Biemer PP, Lyberg, Lars E. Introduction to survey quality. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2003.

63. Olson K, Smyth JD, Ganshert A. The effects of respondent and question characteristics on respondent answering behaviors in telephone interviews. J Surv Stat Methodol [Internet]. 2019;7(2):275–308. Available from: https://academic.oup.com/jssam/article/7/2/275/4968033

64. Goldman SL, Nagel RN, Preiss K, Warnecke HJ. Respondents and quality of survey data. Econ Polit Wkly [Internet]. 2004 [cited 2021 Sep 16];39(7):659–63. Available from: https://www.jstor.org/stable/4414640

65. Knäuper B, Carrière K, Chamandy M, Xu Z, Schwarz N, Rosen NO. How aging affects self-reports [Internet]. Vol. 13, European Journal of Ageing. Springer; 2016 [cited 2021 Sep 16]. p. 185–93. Available from: /pmc/articles/PMC5550601/

66. Ning M, Zhang Q, Yang M. Comparison of self-reported and biomedical data on hypertension and diabetes: findings from the China Health and Retirement Longitudinal Study (CHARLS). BMJ Open. 2016;6(1):e009836.

67. Molenaar EA, Ameijden EJC Van, Grobbee DE, Numans ME. Comparison of routine care self-reported and biometrical data on hypertension and diabetes: Results of the Utrecht Health Project. Eur J Public Health. 2007;17(2):199–205.

68. Peterson KL, Jacobs JP, Allender S, Alston LV, Nichols M. Characterising the extent of misreporting of high blood pressure, high cholesterol, and diabetes using the Australian Health Survey. BMC Public Health. 2016;16(1):1–11.

69. Choi A, Cawley J. Health disparities across education: The role of differential reporting error. Health Econ. 2018 Dec 6;27(3).

70. Fowler F, Mangione T. Standardized Survey Interviewing [Internet]. Standardized Survey Interviewing. 2011 [cited 2021 Sep 16]. 151 p. Available from: https://books.google.com/books/about/Standardized_Survey_Interviewing.html?hl=pt-PT&id=gYyD-WUs35oC

71. Kreuter F. Improving surveys with paradata: analytic uses of process information. Kreuter F, editor. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2013.

72. Greenland S. Statistical uncertainty due to misclassification: Implications for validation substudies. J Clin Epidemiol. 1988 Jan 1;41(12):1167–74.

73. Yi SS, Johns M, Lim S. Use of Regional Data to Validate and Recalibrate Self-reported Hypertension: Highlighting Differences in Immigrant Groups in New York City. J Immigr Minor Heal. 2016;18(1):202–9.

74. Lumley T, Lumley R. Complex Surveys - A Guide to Analysis Using R. Wiley, editor. New Jersey, NJ: John Wiley & Sons; 2010.

75. Heeringa S, West B, Berglund P. Applied survey data analysis. Boca Raton: Chapman & Hall; 2010.

76. Magder LS, Hughes JP. Logistic Regression When the Outcome Is Measured with Uncertainty. Am J Epidemiol Copyr O [Internet]. 1997 [cited 2021 Sep 14];146(2). Available from: https://academic.oup.com/aje/article/146/2/195/165276

77. Lyles RH, Tang L, Superak HM, King CC, Celentano DD, Lo Y, et al. Validation Data-based Adjustments for Outcome Misclassification in Logistic Regression. Epidemiology. 2011 Jul;22(4):589–97.

78. Edwards JK, Cole SR, Chu H, Olshan AF, Richardson DB. Accounting for Outcome Misclassification in Estimates of the Effect of Occupational Asbestos Exposure on Lung Cancer Death. Am J Epidemiol [Internet]. 2014 Mar 1 [cited 2021 Nov 1];179(5):641. Available from: /pmc/articles/PMC3927979/

79. Elliott MR, Raghunathan TE, Schenker N. Combining Estimates from Multiple Surveys. In: Wiley StatsRef: Statistics Reference Online [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2018 [cited 2020 Aug 19]. p. 1–10. Available from: http://doi.wiley.com/10.1002/9781118445112.stat08079

80. Reiter JP. Multiple imputation when records used for imputation are not used or disseminated for analysis. Biometrika [Internet]. 2008 Dec 1 [cited 2021 Nov 1];95(4):933–46. Available from: https://academic.oup.com/biomet/article/95/4/933/262480

81. Carpenter JR, Kenward MG. Multiple Inputation and its Application. 1st ed. John Wiley & Sons, Ltd; 2013.

82. Blackwell M, Honaker J, King G. A Unified Approach to Measurement Error and Missing Data: Overview and Applications. Sociol Methods Res. 2017;46(3):303–41.

83. Seaman SR, White IR, Copas AJ, Li L. Combining Multiple Imputation and Inverse-Probability Weighting. Biometrics [Internet]. 2012 Mar 1 [cited 2021 Nov 1];68(1):129–37. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1541-0420.2011.01666.x

84. Forouzanfar MH, Afshin A, Alexander LT, Anderson HR, Bhutta ZA, Biryukov S, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016 Oct;388(10053):1659–724.

85. Murray CJL, Barber RM, Foreman KJ, Ozgoren AA, Abd-Allah F, Abera SF, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: Quantifying the epidemiological transition. Lancet. 2015;386(10009):2145–91.

86. Wilkins E, Wilson L, Wickramasinghe K, Bhatnagar P, Leal J, Luengo-Fernandez R, et al. European Cardiovascular Disease Statistics 2017 edition. Eur Hear Network, Brussels. 2017;34(39):192.

87. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet [Internet]. 2020 Oct 17 [cited 2021 Nov 1];396(10258):1204–22. Available from: http://www.thelancet.com/article/S0140673620309259/fulltext

88. Instituto Nacional de Estatística. Causas de Morte 2015. 2017th ed. INE, editor. Lisboa: INE;

89. Despacho n.º 6401/2016. Programas de Saúde Prioritários. Diário da República, 2ª série — Nº 94 — 16 maio 2016. 2016;15239.

90. Sociedade Portuguesa de Diabetologia. Diabetes: Factos e Números – O Ano de 2014. Relatório Anual do Observatório Nacional da Diabetes. Lisboa; 2015.

91. Direção Geral da Saúde. Programa Nacional para as Doenças Cérebro-Cardiovasculares. Portugal: Doenças Cérebro-Cardiovasculares em Números 2015. Lisboa; 2015.

92. Gardete-Correia L, Boavida JM, Raposo JF, Mesquita AC, Fona C, Carvalho R, et al. First diabetes prevalence study in Portugal: PREVADIAB study. Diabet Med. 2010 Apr 26;27(8):879–81.

93. Polonia J, Martins L, Pinto F, Nazare J. Prevalence, awareness, treatment and control of hypertension and salt intake in Portugal. J Hypertens. 2014 Jun;32(6):1211–21.

94. European Comission. EUROESTAT. European Health Interview Survey (EHIS wave 3) Methodological manual [Internet]. 2020 [cited 2021 Nov 1]. 1–223 p. Available from: http://ec.europa.eu/eurostat/documents/3859598/8762193/KS-02-18-240-EN-N.pdf/5fa53ed4-4367-41c4-b3f5-260ced9ff2f6

95. Mindell JS, Moody A, Vecino-Ortiz AI, Alfaro T, Frenz P, Scholes S, et al. Comparison of health examination survey methods in Brazil, Chile, Colombia, Mexico, England, Scotland and the USA. Am J Epidemiol. 2017 May 9;

96. Chun H, Kim IH, Min KD. Accuracy of Self-reported Hypertension, Diabetes, and Hypercholesterolemia: Analysis of a Representative Sample of Korean Older Adults. Osong Public Heal Res Perspect. 2016;7(2):108–15.

97. Mozumdar A, Liguori G. Corrective Equations to Self-Reported Height and Weight for Obesity Estimates Among U.S. Adults: NHANES 1999–2008. Res Q Exerc Sport. 2016;87(1):47–58.

98. Gorber SC, Schofield-Hurwitz S, Hardt J, Levasseur G, Tremblay M. The accuracy of self-reported smoking: A systematic review of the relationship between self-reported and cotinine-assessed smoking status. Vol. 11, Nicotine and Tobacco Research. 2009. p. 12–24.

99. Natarajan S, Lipsitz SR, Nietert PJ. Self-report of high cholesterol: Determinants of validity in U.S. adults. Am J Prev Med. 2002;23(1):13–21.

100. Norwood MS, Hughes JP, Amico KR. The validity of self-reported behaviors: methods for estimating underreporting of risk behaviors. Ann Epidemiol. 2016;26(9):612-618.e2.

101. van den Akker M, van Steenkiste B, Krutwagen E, Metsemakers JFM. Disease or no disease? Disagreement on diagnoses between self-reports and medical records of adult patients. Eur J Gen Pract. 2015 Jan 2;21(1):45–51.

102. Hansen H, Schäfer I, Schön G, Riedel-Heller S, Gensichen J, Weyerer S, et al. Agreement between self-reported and general practitioner-reported chronic conditions among multimorbid patients in primary care - results of the MultiCare

Cohort Study. BMC Fam Pract. 2014;15(1):39.

103. Gonçalves VSS, Andrade KRC, Carvalho KMB, Silva MT, Pereira MG, Galvao TF. Accuracy of self-reported hypertension: a systematic review and meta-analysis. J Hypertens [Internet]. 2018 May 1 [cited 2019 Jun 11];36(5):970–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29232280

104. Xie YJ, Ho SC, Liu ZM, Hui SSC. Comparisons of measured and self-reported anthropometric variables and blood pressure in a sample of Hong Kong female nurses. PLoS One. 2014;9(9):10–4.

105. Dave GJ, Bibeau DL, Schulz MR, Aronson RE, Ivanov LL, Black A, et al. Predictors of congruency between self-reported hypertension status and measured blood pressure in the stroke belt. J Am Soc Hypertens. 2013;7(5):370–8.

106. El Fakiri F, Bruijnzeels MA, Hoes AW. No evidence for marked ethnic differences in accuracy of self-reported diabetes, hypertension, and hypercholesterolemia. J Clin Epidemiol. 2007;60(12):1271–9.

107. Tompkins G, Forrest LF, Adams J. Socio-Economic Differences in the Association between Self-Reported and Clinically Present Diabetes and Hypertension: Secondary Analysis of a Population-Based Cross-Sectional Study. PLoS One. 2015;10(10):e0139928.

108. Conti S, Farchi G, Minelli G, Buiatti E, Balzi D, Arniani S, et al. Health Observation Compared to Health Reporting: Findings from a Pilot Study in Florence. Ann Epidemiol. 2007;17(12):999–1003.

109. Barron S, Balanda K, Hughes J, Fahy L. National and subnational hypertension prevalence estimates for the Republic of Ireland: better outcome and risk factor data are needed to produce better prevalence estimates. BMC Public Health. 2014;14(1):24.

110. Mosca I, Bhuachalla BN, Kenny RA. Explaining significant differences in subjective and objective measures of cardiovascular health: evidence for the socioeconomic gradient in a population-based study. BMC Cardiovasc Disord. 2013;13(1):64.

111. Mosca I, Kenny RA. Exploring differences in prevalence of diagnosed, measured and undiagnosed hypertension: the case of Ireland and the United States of America. Int J Public Health. 2014;59(5):759–67.

112. Pereira M, Azevedo A, Barros H. Determinants of awareness, treatment and

control of hypertension in a Portuguese population. Rev Port Cardiol [Internet]. 2010 [cited 2021 Oct 30];29(12):1779–92. Available from: https://pubmed.ncbi.nlm.nih.gov/21428134/

113. Macedo ME, Lima MJ, Silva AO, Alcantara P, Ramalhinho V, Carmona J. Prevalence, awareness, treatment and control of hypertension in Portugal: The PAP study. J Hypertens [Internet]. 2005 [cited 2021 Oct 30];23(9):1661–6. Available from: https://pubmed.ncbi.nlm.nih.gov/16093910/

114. Joint National Committee on Detection, Evaluation and T of HBP. The fifth report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure. Arch Intern Med 1993;153154–83.

115. Committee GS of the WMHL. 1993 guidelines for the management of mild hypertension. Memorandum from a World Health Organization/International Society of Hypertension meeting. Hypertension. 1993 Sep;22(3):392–403.

116. Gwynn RC, Garg RK, Kerker BD, Frieden TR, Thorpe LE. Contributions of a Local Health Examination Survey to the Surveillance of Chronic and Infectious Diseases in New York City. 2009;99(1):152–9.

117. Van Eenwyk J, Bensley L, Ossiander E, Krueger K. Comparison of Examination-Based and Self-Reported Risk Factors for Cardiovascular Disease, Washington State, 2006–2007. Prev Chronic Dis. 2012;9:1–12.

118. Direção Geral da Saúde. Abordagem Terapêutica das Dislipidemias no Adulto. Lisboa; 2013 p. 1–17.

119. Gorber S, Tremblay M, Campbell N, Hardt J. The Accuracy of Self-Reported Hypertension: A Systematic Review and Meta-Analysis. Curr Hypertens Rev. 2008 Jan 30;4(1):36–62.

120. Goldman N, Lin IF, Weinstein M, Lin YH. Evaluating the quality of self-reports of hypertension and diabetes. J Clin Epidemiol. 2003;56(2):148–54.

121. Man REK, Gan AHW, Fenwick EK, Gan ATL, Gupta P, Sabanayagam C, et al. Prevalence, determinants and association of unawareness of diabetes, hypertension and hypercholesterolemia with poor disease control in a multi-ethnic Asian population without cardiovascular disease. Popul Health Metr [Internet]. 2019 Dec 5 [cited 2021 Nov 2];17(1). Available from: /pmc/articles/PMC6896313/

122. Aguiar P, Nunes B. [{Odds} {Ratio}: review about the meaning of an epidemiological measure]. Acta Med Port. 2013;26(5):505–10.

123. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol. 2003 Jun 20;3(1):21.

124. McNutt L-A, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. Am J Epidemiol. 2003 May 15;157(10):940–3.

125. Wagner KJP, Boing AF, Subramanian S V, Höfelmann DA, D'Orsi E. Effects of neighborhood socioeconomic status on blood pressure in older adults. Rev Saude Publica. 2016;50(0):78.

126. Raftopoulou A. Geographic determinants of individual obesity risk in Spain: A multilevel approach. Econ Hum Biol. 2017;24:185–93.

127. Mandal B, Chern WS. A Multilevel Approach to Model Obesity and Overweight in the U.S. Int J Appl Econ. 2011;8(2):1–17.

128. Maier W, Holle R, Hunger M, Peters A, Meisinger C, Greiser KH, et al. The impact of regional deprivation and individual socio-economic status on the prevalence of Type 2 diabetes in Germany. A pooled analysis of five population-based studies. Diabet Med. 2013 Mar;30(3):e78–86.

129. Johnston DW, Propper C, Shields MA. Comparing Subjective and Objective Measures of Health : Evidence from Hypertension for the Income / Health Gradient Comparing Subjective and Objective Measures of Health : Evidence from Hypertension for the Income / Health Gradient. 2009;2737(2737):540–52.

130. Burgard SA, Chen P V. Challenges of health measurement in studies of health disparities. Soc Sci Med. 2014;106:143–50.

131. Mackenbach JP, Looman CWN, Van Der Meer JBW. Differences in the misreporting of chronic conditions, by level of education: The effect on inequalities in prevalence rates. Am J Public Health. 1996 May;86(5):706–11.

132. Vellakkal S, Millett C, Basu S, Khan Z, Aitsi-Selmi A, Stuckler D, et al. Are estimates of socioeconomic inequalities in chronic disease artefactually narrowed by self-reported measures of prevalence in low-income and middle-income countries? Findings from the WHO-SAGE survey. J Epidemiol Community Health. 2015 Mar;69(3):218–25.

133. Mindell JS, Giampaoli S, Goesswald A, Kamtsiuris P, Mann C, Männistö S, et al. Sample selection, recruitment and participation rates in health examination surveys in Europe--experience from seven national surveys. BMC Med Res

Methodol. 2015;15:78.

134. Nunes B, Barreto M, Gil AP, Kislaya I, Namorado S, Antunes L, et al. The first Portuguese National Health Examination Survey (2015): design, planning and implementation. Eur J Public Health. 2016 Nov;

135. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009 Apr;42(2):377–81.

136. Tolonen H, Koponen P, Aromaa A, Conti S, Liv SG, Kanieff M, et al. Review of Health Examination Surveys in Europe. Training. 2008;

137. Tolonen H. EHES Manual. Part A. Planning and preparation of the survey. Vol. 2013. 2013.

138. Tolonen H. EHES Manual. Part B. Fieldwork procedures. 2013;

139. Tolonen H, Dobson A, Kulathinal S, WHO MONICA Project. Assessing the quality of risk factor survey data: lessons from the WHO MONICA Project. Eur J Cardiovasc Prev Rehabil. 2006 Feb;13(1):104–14.

140. Ferrario M, Kuulasmaa K, Grafnetter D, Moltchanov V. Quality assessment of total cholesterol measurements in the WHO MONICA Project [Internet]. National Institute for Health and Welfare, Finland, on behalf of the World Health Organization and the WHO MONICA Project investigators; 1999 Nov [cited 2019 Feb 13]. Available from: https://thl.fi/publications/monica/tchol/tcholqa.htm

141. Tolonen H, Ferrario M, Minoja M, WHO MONICA Project. Quality Assessment of Data on Awareness and Treatment of High Cholesterol in the WHO MONICA Project. 1999 Nov 5 [cited 2019 Feb 13]; Available from: http://www.thl.fi/publications/monica/hich/hchdrug.htm

142. Instituto Nacional de Estatística. Inquerito Nacional de Saúde 2014 [Internet]. INE. Lisboa: INE; 2016. 310 p. Available from: www.ine.pt

143. Tolonen H. EHES Manual Part C. European level collaboration. 2013.

144. Fifth Joint Task Force of the European Society of Cardiology, European Association of Echocardiography, European Association of Percutaneous Cardiovascular Interventions, European Heart Rhythm Association, Heart Failure Association, European Association for Cardiovascular Prevention & Rehabilitation, et al. European Guidelines on cardiovascular disease prevention

in clinical practice (version 2012). Eur J Prev Cardiol. 2012 Aug 4;19(4):585–667.

145. Reiner Z, Catapano AL, De Backer G, Graham I, Taskinen M-R, Wiklund O, et al. ESC/EAS Guidelines for the management of dyslipidaemias: The Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS). Eur Heart J. 2011 Jul 2;32(14):1769–818.

146. Eurostat. European Health Interview Survey wave 2 methodological manual. 2013. 1–202 p.

147. Eurostat. Statistical Requirements Compedium.

148. Lumley R. Complex surveys: a guide to analysis using R. Hoboken, NJ: John Wiley & Sons; 2010.

149. Lehtonen R, Pahkinen E. Practical Methods for Design and Analysis of Complex Surveys. 2nd ed. Jonh Wiley & Sons; 2004.

150. Departamento de Epidemiologia. I Inquérito Nacional de Saúde com Exame Físico (INSEF I): Relatório Metodológico. Lisboa; 2016.

151. J. N. K. Rao, A. J. Scott. On Chi-Squared Tests for Multiway Contingency Tables with Cell Proportions Estimated from Survey Data. Ann Stat [Internet]. 1984 [cited 2017 Jul 5];12(1):46–60. Available from: https://www.jstor.org/stable/2241033

152. Rao JNK, Scott AJ. On Simple Adjustments to Chi-Square Tests with Sample Survey Data. Ann Stat. 1987 Mar;15(1):385–97.

153. Ernstsen L, Strand BH, Nilsen SM, Espnes GA, Krokstad S. Trends in absolute and relative educational inequalities in four modifiable ischaemic heart disease risk factors: repeated cross-sectional surveys from the Nord-Trøndelag Health Study (HUNT) 1984–2008. BMC Public Health. 2012 Apr 3;12(1):266.

154. Mackenbach JP, Kunst AE. Measuring the magnitude of socio-economic inequalities in health: An overview of available measures illustrated with two examples from Europe. Soc Sci Med. 1997;44(6):757–71.

155. Departamento de Epidemiologia; Departamento de Promoção da Saúde e Doenças Crónicas. Inquérito de Saúde com Exame Físico. relatório de estudo realizado em São Brás de Alportel. Instituto Nacional de Saúde Doutor Ricardo Jorge, IP; 2012.

156. StataCorp. Stata Statistical Software: Release 15. College Station, TX: StataCorp LP; 2017.

157. Cortez-Dias N, Robalo Martins S, Belo A, Fiúza M. Characterization of lipid profile in primary health care users in Portugal. Rev Port Cardiol (English Ed [Internet]. 2013 Dec 1 [cited 2021 Sep 20];32(12):987–96. Available from: http://www.revportcardiol.org/en-characterization-lipid-profile-in-primary-articulo-S217420491300202X

158. da Silva PM, Aguiar C, Morais J. Suboptimal lipid levels in clinical practice among Portuguese adults with dyslipidemia under lipid-lowering therapy: Data from the DISGEN-LIPID study. Rev Port Cardiol. 2019 Aug 1;38(8):559–69.

159. Pereira M, Carreira H, Vales C, Rocha V, Azevedo A, Lunet N. Trends in hypertension prevalence (1990–2005) and mean blood pressure (1975–2005) in Portugal: a systematic review. http://dx.doi.org/103109/080370512012666380 [Internet]. 2012 Aug [cited 2021 Oct 30];21(4):220–6. Available from: https://www.tandfonline.com/doi/abs/10.3109/08037051.2012.666380

160. Mackenbach JP, Cavelaars AEJM, Kunst AE, Groenhof F, Andersen O, Borgan JK, et al. Socioeconomic inequalities in cardiovascular disease mortality. An international study. Eur Heart J. 2000 Jul 15;21(14):1141–51.

161. Schifeling T, Reiter JP, DeYoreo M. Data Fusion for Correcting Measurement Errors. 2016;1–31.

162. Groves RM, John Wiley & Sons. Survey errors and survey costs. Wiley-Interscience; 2004. 590 p.

163. Utentes Inscritos em Cuidados de Saúde Primários — Transparência [Internet]. [cited 2021 Nov 1]. Available from: https://transparencia.sns.gov.pt/explore/dataset/utentes-inscritos-em-cuidados-de-saude-primarios/information/?disjunctive.ars&disjunctive.aces&sort=periodo

164. Assembleia da República. Lei de Bases da Saúde [Internet]. Vol. 1ª série, Diário da República. 2019 [cited 2021 Nov 1]. p. 55–66. Available from: https://dre.pt/dre/detalhe/lei/95-2019-124417108

165. Presidência do Conselho de Ministros. Decreto-Lei n.º 96/2020. 2021. p. Diário da República n.º 215/2020, Série I de 2020-.

Appendix

Appendix A: Summary of research on accuracy of self-reported hypertension, hypercholesterolemia

First author, year	Sample size	Population	Age Range	Cut-off for hypertension	Statistical methods
Taylor, 2010	1537	General population, Adelaide, Australia	18+	140/90	Absolute difference, Sensitivity, Specificity, PPV, NPV, Cohen's kappa
Ning, 2016	13610	General population, China	45+	140/90, corrected for medication	Sensitivity, Specificity, Cohen's kappa, logistic regression
Huerta, 2009	1556	General population, Murcia, Spain	20+	140/90, corrected for medication	Sensitivity, Specificity, PPV, NPV, Cohen's kappa, Logistic regression
Molenaar, 2006	4950	General population, Leidsche Rijn, Utrecht, Netherlands	18+	140/90 for age <60 years; 160/90 for age 60+	Relative difference, Sensitivity, Specificity, PPV, NPV, Cohen's kappa, Logistic regression
VanEenwyk, 2012	672	General population, Washington State, USA	25+	140/90, corrected for medication	Poisson regression
Gwynn, 2009	9585	General population, New York City, USA	20+	140/90, corrected for medication	% of undiagnosed disease
Tolonen, 2014	4127	General population, locally in 12 European counties	25-64	140/90, corrected for medication	Absolute difference, Sensitivity, Specificity, PPV, NPV
Vargas, 1997	8409	General population, USA, NHANES	25+	140/90, corrected for medication	Sensitivity, Specificity, PPV, NPV, chi-square test
Dave, 2013	17382	Stroke belt population, counties of North Carolina, USA; medicated excluded	18+	140/90	Sensitivity, Specificity, PPV, NPV, Cohen's kappa, Logistic regression
Xie, 2014	216	Female nurses, Hong Kong	35-65	140/90	Sensitivity, Specificity, Cohen's kappa, McNemar test, Bland-Altman plot
Chun, 2016	7270	General population, Korea, KNHANES	50+	140/90, corrected for medication	Sensitivity, Specificity, Cohen's kappa, Logistic regression
Tenkorang,2015	13561, 5071, 10870,	General population, Ghana, China, India, South Africa, and Russia; SAGE project	50+	140/90, corrected for medication	Sensitivity, Specificity, Multinomial logit model for concordance vs. 2 types of error

	4081, 3908				
Conti, 2007	625	General population, Florence, Italy	35-74	160/95, corrected for medication	Sensitivity, Specificity, Cohen's kappa
Atwood, 2013		General population, Canada	20-79	140/90, corrected for medication	prevalence of HTA from different sources
Tsai, 2012	1021	General population, Taiwan	54+	both 140/90 e 160/95, corrected for medication	Sensitivity, Specificity, PPV, NPV, Cohen's kappa, Logistic regression
Barron, 2014	1207	General population, Ireland	45+	140/90, corrected for medication	prevalence of undiagnosed disease
Mosca, 2013	4179	General population, Ireland	50+	140/90, corrected for medication	Logistic regression separately for self-reported and measured
Goldman, 2003	1023	General population, Taiwan	54+	140/90, corrected for medication	Sensitivity, Specificity, Logistic regression
Fakiri, 2007	430	Patients at high risk of cardiovascular disease, Rotterdam and The Hague	30+	160/95	Sensitivity, Specificity, Cohen's kappa, Logistic regression for overall concordance
Tompkins, 2015	3131	General population, Scotland, Scottish Health Survey	18+	140/90, corrected for medication	Sensitivity, Specificity, PPV NPV, Logistic regression
Gracuani, 2013 (2008-10)	11408	Spanish general population	18+	140/90	Awareness, proportion of lack of awareness by age group
Man, 2019	5386	Singapure, general population	40-80	140/90, corrected for medication	Awareness and lack of awareness, logistic regression for lack of awareness
Gee, 2012	3473	Canada, general population	20-79	140/90, corrected for medication	Awareness logistic regression
Vallee, 2020	2015	France, General population	18-74	140/90, corrected for medication	Awareness by age group and sex

First author, year	Sample size	Population	Age Range	Cut-off for Hypercholesterolemia	Statistical methods
Taylor, 2010	1 537	General population in Adelaide, Austrália	18+	TC >=5.5mmol/L	absolte difference, sensitivity, specificity, PPV, NPV, Kappa
Huerta, 2009	1556	General population, Murcia, Spain	20+	TC ≥200 mg/dL, triglycerides ≥200mg/dL, corrected for medication	sensitivity, specificity, PPV, NPV, Kappa, X^2, LR
Van Eenwyk, 2012	672	General population, Washington State, USA	25+	fasting blood samples, total cholesterol of 200–239 mg/dL; LDL cholesterol of 100–159 mg/dL; VLDL cholesterol of 30–39 mg/dL; HDL cholesterol of 40–44 mg/dL for men or 50–54 mg/dL for women; or triglycerides of 150–199 mg/dL; corrected for medication	Poisson regression
Gwynn, 2009	9585	General population, New York City, USA	20+	TC ≥240 mg/dL, corrected for medication	% of undiagnosed disease
Tolonen, 2014	4127	General population, localy in 12 European counties	25-64	TC ≥5.0 mmol/l, corrected for medication	absolte difference, sensitivity, specificity, PPV, NPV
Chun, 2016	7 270	General population, Korea, KNHANES	50+	TC≥240 mg/mL, corrected for medication	sensitivity, specificity,Kappa, LR
Huang, 2007	24 069	Female, health professionals with no prior CVD, USA	45+	14 categories	logistic regression
Mosca, 2013	4 179	General population,Ireland	50+	TC ≥5.2 mmol/L, corrected for medication	Binary logistic regression separatly for self-reported and measured
Fakiri, 2007	430	Patients at high risk of cardiovascular disease, Rotterdam and The Hague	30+	TC >6.5 mmol/L, corrected for medication	sensitivity, specificity, kappa, LR for overall concordance
Natarajan, 2002	8236	General population,USA, NHANES	21+	TC 200 mg/dL, 240mg/dL, e 300mg/dL,	sensitivity, specificity,PPV, NPV, LR
Petterson, 2016		General, Australia		TC >=5.5mmol/L	overreporting, underreporting
Ahluwalia, 2009	733	female, low-income, WV, USA	40-64	nonfasting TC ≥200 mg/dL and TC ≥240 mg/dL corrected for medication	sensitivity, specificity, kappa, ppv, npv

Appendix B List of variables collected by INSEF 2015

Variable	Description	Type	Values	
ID	Identification code	Numeric		
PSU	Primary sampling unit	Categorical		
STRATA	Sampling stratum	Categorical		
SAMPLING_WEIGHT	Sampling weight	Numeric		
URB_TYPE	Degree of urbanization	Categorical	1	Rural
			2	Urban
REGION	Health Region	Categorical	1	Norte
			2	Centro
			3	LVT
			4	Alentejo
			5	Algarve
			6	RA Madeira
			7	RA Açores
SEX	Sex	Categorical	0	Female
			1	Male
AGE_G2	Age group (5-year)	Categorical	1	25-29
			2	30-34
			3	35-39
			4	40-44
			5	45-49
			6	50-54
			7	55-59
			8	60-64
			9	65-69
			10	70-74
HATLEVEL	Highest level of education completed	Categorical	-2	Refusal
			-1	Unknown
			0	No formal education
			1	Basic 1st cycle
			2	Basic 2nd cycle
			3	Basic 3rd cycle
			4	Secondary
			5	Post-secondary
			6	Higher
			7	Post graduate (Mcs/PhD)
OCUP	Employment status	Categorical	-1	Unknown
			1	Employed
			2	Unemployed

			3	Student
			4	Retired
			5	Disabled
			7	Housewife
			8	Other
HHNBPERS	Household size	Categorical	1	1 person
			2	2 persons
			3	3 persons
			4	4 or more persons
NCHILD	Number of children <14 years in the household	Categorical	0	0
			1	1
			2	2 or more
HS1	Self-perceived health	Categorical	-2	Refusal
			-1	Unknown
			1	Very good
			2	Good
			3	Fair
			4	Bad
			5	Very bad
CD	Longstanding illness/health problem (6 months or more)	Categorical	-2	Refusal
			-1	Unknown
			0	No
			1	Yes
			1	Yes
CD3A	Hypertension diagnosed by a medical doctor	Categorical	-3	NA
			-2	Refusal
			-1	Unknown
			0	No
			1	Yes
CD3S	Hypercholesterolemia diagnosed by a medical doctor	Categorical	-3	NA
			-2	Refusal
			-1	Unknown
			0	No
			1	Yes
CD4A	Any first degree relative affected by hypertension	Categorical	-3	NA
			-2	Refusal
			-1	Unknown
			0	No
			1	Yes
CD4S	Any first degree relative affected by hypercholesterolemia	Categorical	-3	NA
			-2	Refusal

			-1	Unknown
			0	No
			1	Yes
MED1	Use of prescribed medicines in the past 2 weeks	Categorical	-2	Refusal
			-1	Unknown
			0	No
			1	Yes
MED1A	Prescribed medication for hypertension	Categorical	-2	Refusal
			-1	Unknown
			0	No
			1	Yes
MED1S	Prescribed medication for hypercholesterolemia	Categorical	-2	Refusal
			-1	Unknown
			0	No
			1	Yes
BT2	Last cholesterol test	Categorical	-2	Refusal
			-1	Unknown
			1	Less than 3 months
			2	3-5 months
			3	6-11 months
			4	12 months or more
			5	Never
BP_M	Last blood pressure measurement by a health professional	Categorical	-2	Refusal
			-1	Unknown
			1	Less than 3 months
			2	3-5 months
			3	6-11 months
			4	12 months or more
			5	Never
HOSPIT1	Hospitalization in the last 12 months	Categorical	-2	Refusal
			-1	Unknown
			0	No
			1	Yes
GP1	Last consultation with a general practitioner (GP) or family doctor in a health centre	Categorical	-2	Refusal
			-1	Unknown
			1	Less than 12 months ago
			2	12 months ago or longer
			3	Never
MS1	Last medical or surgical specialist consultation	Categorical	-2	Refusal
			-1	Unknown
			1	Less than 12 months ago

			2	12 months ago or longer
			3	Never
INC	Categories of income per equivalent adult (OECD-modified scale)	Categorical	-2	Refusal
			-1	Unknown
			1	Low
			2	Medium-low
			3	Medium
			4	Medium-high
			5	High
INC2	Household can afford an unexpected expense of 434 euros without asking for a loan	Categorical	-3	NA
			-2	Refusal
			-1	Unknown
			0	No
			1	Yes
CTOT	Amount of total cholesterol (in milligrams per decilitre)	Numeric	-1	Unknown
HDL	Amount of high-density lipoprotein cholesterol (in milligrams per decilitre)	Numeric	-1	Unknown
LDL	Amount of low-density lipoprotein cholesterol (in milligrams per decilitre)	Numeric	-1	Unknown
PREG	Pregnancy	Categorical	-3	NA
			-2	Refusal
			-1	Unknown
			0	No
			1	Yes
BP_SYST	Systolic blood pressure (mmHg)	Numeric	-1	Unknown
BP_DIAST	Diastolic blood pressure (mmHg)	Numeric	-1	Unknown
BMI	Body mass index (kg/m2)	Numeric	-1	Unknown