

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}

^aDepartamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisbon, Portugal; ^bCentro de Investigação em Saúde Pública, Escola Nacional de Saúde Pública, Universidade Nova de Lisboa, Lisbon, Portugal;

^cDepartment of Health and Inequality, Division of Mental and Physical Health, Norwegian Institute of Public Health, Oslo, Norway; ^dInterdisciplinary Centre of Social Sciences (NOVA FCSH), Lisbon, Portugal; ^eDepartment 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.

© 2020 The Author(s) Published by S. Karger AG, Basel
on behalf of NOVA National School of Public Health

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

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 (IN-SEF) 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.

© 2020 The Author(s). Published by S. Karger AG, Basel on behalf of NOVA National School of Public Health

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 de-

tailed 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 to 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 haemo-

lysed 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 Verification of the last digit distribution for height	Pilot and PSU for all the regions (PSU1 through PSU7) Pilot and PSU for all the regions (PSU1 through PSU7)	Minimum duration of 8 min –
Blood collection and processing	Proportion of haemolysed serum samples	PSU for all the regions (PSU1 through PSU7)	Below 5%
Health questionnaire	Average interview time Average rate of items missing for selected questionnaire areas	PSUs for all the regions (PSU1 through PSU7) Overall	– Below 5%
PSU, primary sampling unit.			

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

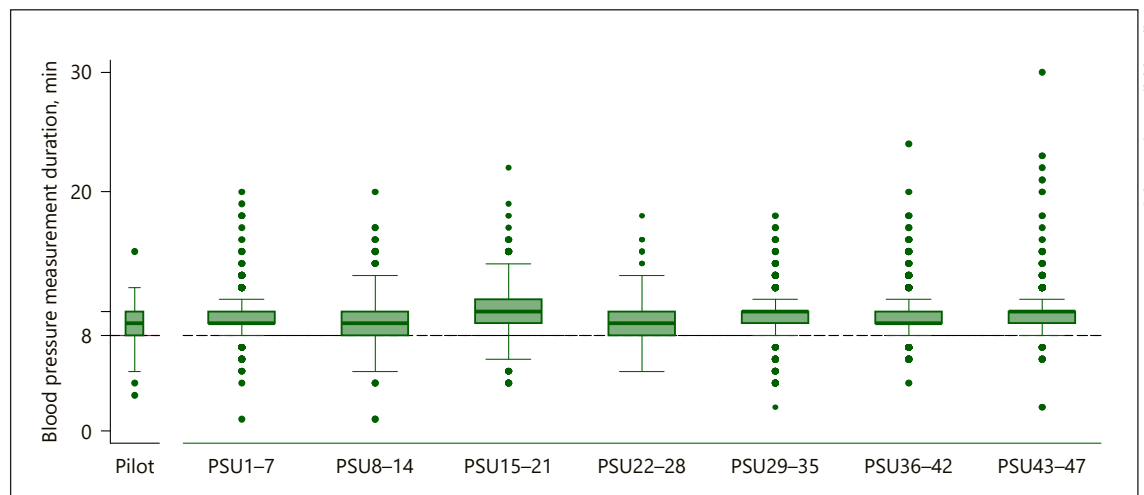
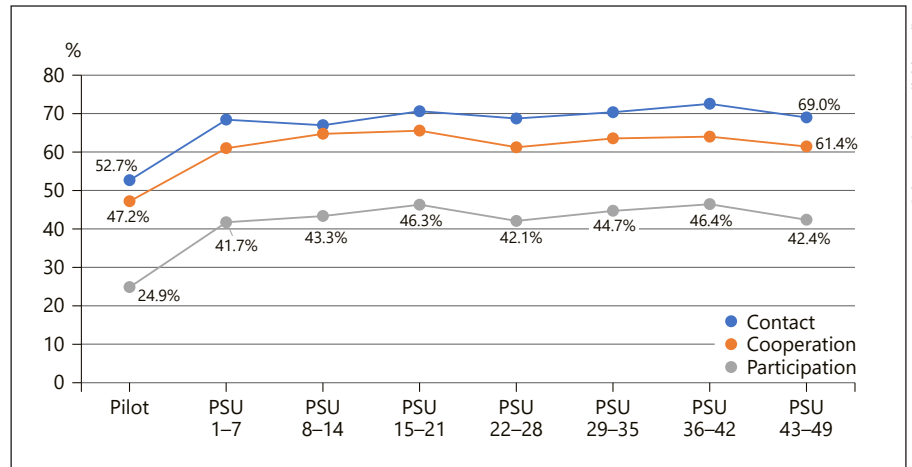


Fig. 2. Distribution of blood pressure measurement duration (min) over fieldwork. PSU, primary sampling unit.

(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

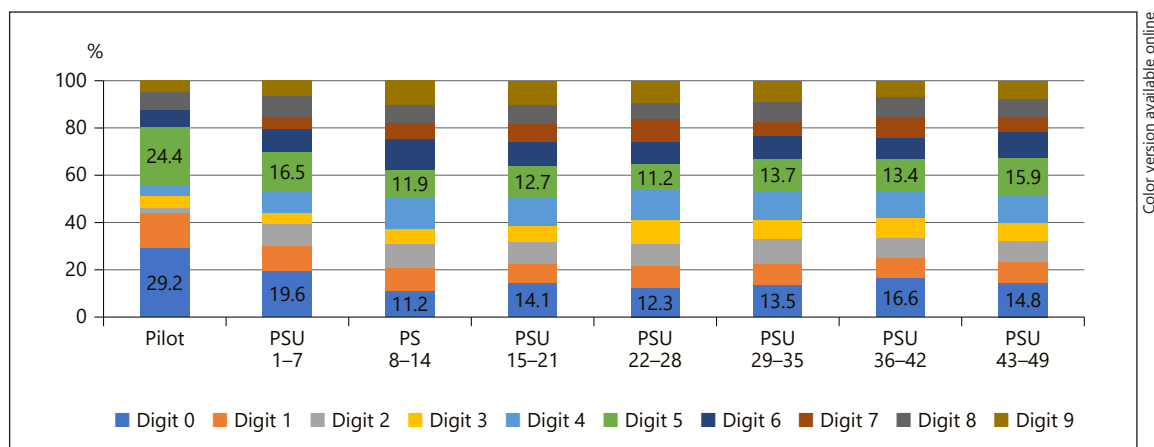


Fig. 3. Last digit distribution for height measurement throughout the fieldwork. PSU, primary sampling unit.

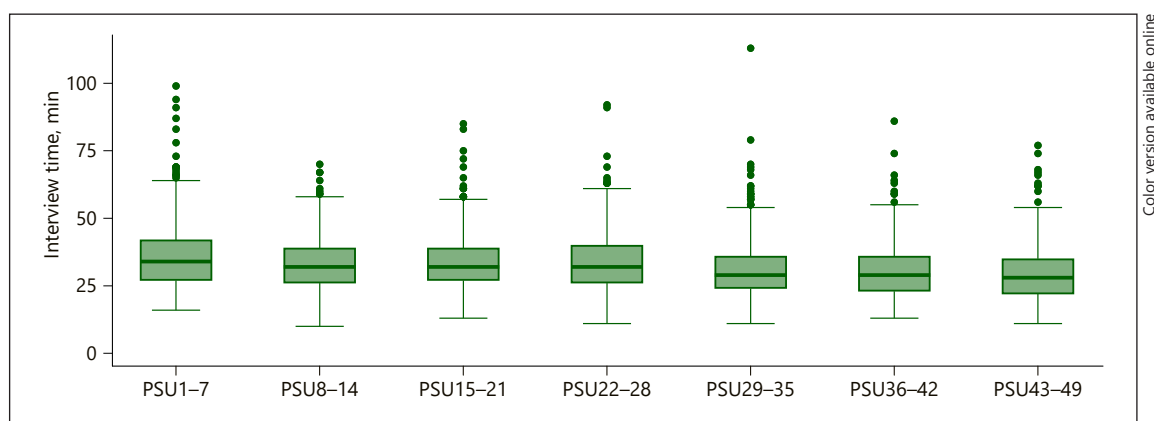


Fig. 4. Average interview time (min) throughout the fieldwork. PSU, primary sampling unit.

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” par-

ticipants 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 previ-

ously 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>.