

RESEARCH ARTICLE

Open Access

Nurse effects on measurement error in household biosocial surveys



Alexandru Cernat^{1*}  and Joseph W. Sakshaug^{2,3,4}

Abstract

Background: Biosocial survey data are in high demand, yet little is known about the measurement quality of health measures collected by nurses in respondents' homes. Our objective was to analyze the degree to which nurses influence measurement in anthropometric and physical performance indicators collected from respondents in two nationally-representative UK biosocial surveys.

Methods: The English Longitudinal Survey of Ageing and the UK Household Longitudinal Study – Understanding Society were used to analyze fourteen anthropometric and physical performance measures covering weight, height, pulse, grip strength, and lung capacity. Cross-classified multilevel models were used to estimate “nurse effects” on measurement error.

Results: Overall, there is a medium effect of nurses on measurement. Across all measures collected in both studies, nurses explain around 13% of all measurement variation. Variation in specific measures range between approximately 2 and 25%. Grip strength and lung capacity are more heavily influenced by nurses than are height, weight, and pulse. Lastly, nurse characteristics explain only a very small proportion of nurse measurement variation.

Conclusion: Objective health measures collected by nurses in household biosocial surveys are susceptible to non-trivial amounts of measurement variation. Nurse ID numbers should be regularly included in biosocial data releases to allow researchers to account for this unnecessary source of variation. Further, researchers are advised to conduct sensitivity analyses using control variables that account for nurse variation to confirm whether their substantive findings are influenced by nurse measurement effects.

Keywords: Biosocial surveys, Measurement error, Nurse effects, Anthropometric measures

Background

The collection of objective health measures (or “biomeasures”) in population-based social surveys has become increasingly prevalent over the years [1]. At present, numerous surveys collect anthropometric measures (e.g. height, weight, waist and hip circumference, blood pressure), physical performance measures (e.g. grip strength, lung function), and biological specimens (e.g. blood, urine) alongside traditional survey measures [2–5]. These so-called “biosocial surveys” provide researchers with the means to enhance their understanding of the complex interrelationships between the social environment and health outcomes in the population [6–8].

There are at least three models of collecting biomeasures in social surveys. One model consists of a center-based assessment, where respondents present themselves at a medical facility (hospital, clinic, health center) [9, 10]. This model allows for a wide range of biomeasures to be collected with sophisticated equipment administered by a licensed medical professional. However, this is the costliest model and impractical for older, more vulnerable populations [11]. A less expensive model is to train lay interviewers to administer the biomeasures inside respondents' homes [12–16]. This model conveniently allows the administration of both the traditional interview and collection of biomeasures in a single visit. However, the range of biomeasures that can be collected through this model is more limited as some countries require certain specimens (e.g. whole blood) to be collected by a licensed medical professional. A compromise

* Correspondence: alexandru.cernat@manchester.ac.uk

¹University of Manchester, Humanities Bridgford Street, Manchester M13 9PL, UK

Full list of author information is available at the end of the article



on the center-based and lay interviewer models is to send licensed nurses to respondents' homes at some point after the traditional interview takes place [3, 17]. This model allows for the collection of a broader range of biomeasures compared to the lay interviewer model and at a significantly lower cost compared to the centre-based model.

Any model that deploys actors to collect biomeasures in nonclinical and nonstandardized settings poses challenges that can potentially affect the quality of the collected measures [18]. Quality assessments of biomeasures collected by lay interviewers have been documented [4, 15, 19, 20], but quality assessments for nurses who are deployed to respondents' homes in social surveys are still lacking. A large body of work indicates that nurses are prone to committing measurement errors in clinical settings [21–29]. Imprecision of measurement (or between-observer variation) is the most commonly cited source of measurement error in nurse-led assessments [28, 29]. Biomeasures found to be susceptible to nurse measurement error, include height, weight, and waist and hip circumference [28], blood pressure [21, 22, 25], and skin fold measurements [28]. Multiple factors have been suggested as possible causes of nurse measurement error for these anthropometric measures, such as incorrect cuff size, variation in training, digit rounding, inadequate knowledge of appropriate measurement techniques, and the use of multiple nurses [24, 28, 30–35].

Suggested guidelines for minimizing nurse measurement error include adequate training, ongoing quality control, a controlled working environment free of impediments, double measurement of a subsample to assess the extent of imprecision, and minimizing the number of nurses used within a given study [28, 34, 36–38]. However, these guidelines are incongruent with the practical realities of most large-scale biosocial surveys. First, freelance nurses are decentralized and geographically dispersed which makes standardized training and ongoing oversight of their technique rarely feasible. Second, the working environment (i.e. respondents' homes) can vary considerably with respect to space, lighting, and other factors that may influence measurement. Third, double measurement is impractical as it places undue burden on respondents and may harm participation rates in follow-up waves of a longitudinal study [39]. Lastly, many nurses are needed to cover large geographical areas common in nationally-representative biosocial surveys. Deploying a large number of nurses increases the possible magnitude of measurement error, even when only small differences in technique occur between nurses over time [28].

Given the high demand for biosocial survey data, it is important to understand their impact on the measures

collected. Only after that can the quality of these measurements be determined. While some biomeasures are relatively simple to administer and require modest amounts of training (e.g. anthropometric measures), there are others which are more complex and require more extensive training (e.g. physical performance measures). Assessing the degree of nurse measurement error across multiple biomeasures with different administration difficulties is therefore needed to determine where larger measurement errors occur and where improvements in measurement technique are most needed.

In this article, we examine the extent of nurse effects in two longitudinal household biosocial surveys: the English Longitudinal Study of Ageing (ELSA) and Understanding Society – the UK Household Longitudinal Study (US). Nurse effects are defined as variability in the measurements at the nurse level. This form of nurse measurement error introduces non-zero correlations among the measurements collected by a nurse, which can inflate the variance of descriptive estimates. Using a cross-classified multilevel model that separates nurse and area effects, we estimate nurse effects for a host of anthropometric and physical performance measures collected over multiple waves of each study. Lastly, we make use of background data on the nurses themselves to determine whether their age and level of experience are factors that explain nurse measurement error and should be accounted for in analyses of biosocial survey data.

In short, the following three research questions are addressed:

- 1) To what extent do nurses contribute to measurement error in biomeasure collection?
- 2) Are nurse measurement error effects consistent across surveys and over time?
- 3) Do nurse characteristics explain a significant amount of nurse measurement error?

Methods

Data sources

The UK Household Longitudinal Study (UKHLS) is a panel survey representative of the UK population [40]. It initially started in 1992 under the name British Household Panel Survey (BHPS). In 2009 a new version of the survey was implemented with a sample size of 40,000 households in the UK under the name Understanding Society (US). In wave 2 of the Understanding Society (USW2), a random 80% of the sample was selected for a nurse visit. Trained nurses visited respondents in their homes around 6 months after the main interview. The nurse visit collected data on height, weight, pulse, grip strength and lung capacity after receiving verbal consent. Additionally, after receiving written consent nurses also

collected whole blood from the vein, but this component is not part of the present investigation. In wave 3 of Understanding Society (or wave 19 of the BHPS; BHPSW19), all eligible BHPS members were selected for a nurse visit. Similar procedures were followed as in the previous wave.

In US wave 1 the individual response rate was 81.8% while in waves 2 and 3 they were 59.4 and 61.3%, respectively [41]. Response rates for the nurse visit (among eligible individuals) were 58.6% for USW2 and 57% for BHPSW19 [42].

The English Longitudinal Study of Ageing (ELSA) is a longitudinal study that collects information every 2 years from a representative sample of residents in England who are 50 years of age and older [17]. The sample is based on respondents from the Health Survey for England. In ELSA waves 2 (ELSAW2), 4 (ELSAW4), and 6 (ELSAW6), nurses visited eligible respondents and collected the same anthropometric and physical performance measures as in the US/BHPS studies, as well as blood from the vein.

For ELSA waves 2, 4 and 6 the individual response rates were 82, 71, and 76%, respectively [43–46], while response rates for the nurse visits were 87.3, 85.7 and 84.3%, respectively, among eligible individuals [47].

Outcomes

To investigate the impact of nurse effects on measurement error we consider five different outcome measures: height, weight, pulse, grip strength, and lung capacity. These cover the typical anthropometric and physical performance measures collected by nurses. Some of these measures are administered multiple times. For example, pulse is measured three times, grip strength is measured three times for each hand, and lung capacity is measured three times based on the equipment used. Table 1 shows all of the measurements collected. In total, fourteen individual measurements were taken [42,

48]. In the forthcoming analysis, each individual measurement is analyzed separately to evaluate whether nurses have a differential impact on the full range of measurements they collect.

To investigate the influence of nurses on measurement we adopt similar procedures to those used in the interviewer effects literature [49]. The main challenge in this research is separating the effect of the nurse from other possible confounders, especially area effects and respondent characteristics. In the absence of randomized allocation of nurses to respondents, a statistical approach is needed to control for these confounders. To separate nurse and area effects, a cross-classified multi-level model is used with random effects for nurses and areas [50]. Here, areas are defined as Lower Super Output Areas that represent areas of approximately 1500 households.

Respondent characteristics are introduced as control variables. The control variables used for both surveys are: sex, age, having a partner, owning the house, education, overall health, if they have a long-term illness, if they live in London and if they live in the north of UK. Additional variables are included for the models based on the US data: if living in an urban area, household size, and interest in politics. In ELSA, a variable was included on whether the respondent lives alone.

Statistical analyses

The cross-classified multilevel model is defined as:

$$Y_{i(j,k)} = \gamma_0 + \sum \gamma_h x_{i(j,k)} + U_{0j} + U_{0k} + \varepsilon_i$$

where Y , the dependent variable, varies by individual (i), area (j), and nurse (k). This model is explained by an intercept (γ_0) term and h control variables with fixed effects (γ_h). The random effects for area (U_{0j}) and nurse (U_{0k}) are cross-classified. Lastly, ε_i represents the residual, or unexplained variance.

Table 1 Measures collected and analyzed in the US/BHPS and ELSA surveys

Type	Measure	Units	Number of measurements	Name in data
Anthropometric measures	Height	cm	1	height
	Weight	kg	1	weight
	Pulse	Beats per minute	3	pulse
Physical performance	Grip strength	Individual readings for non-dominant hand in kg	3	mmgsn
		Individual readings for dominant hand in kg	3	mmgsd
	Lung capacity	The amount of air that can be blown out in one second, measured in liters	1	htfev
		The speed of air moving out of lungs at the beginning of expiration, measured in liters per second	1	htpev
	The total amount of air that can be forcibly blown out after a full inhalation, measured in liters	1	htfvc	

In the context of the present research, U_{0k} represents nurse effects. It represents the amount of variation that is explained by nurses after controlling for respondent characteristics and area effects. If nurses have no impact on the collection of a biomeasure this should be close to zero. This quantity will answer the first research question. To answer the second question, we inspect how this coefficient varies by the survey, wave of data collection, and biomeasure.

To answer the third research question, the model is expanded to include nurse control variables. Two nurse-level variables were collected in both surveys: nurse age and nurse experience. Comparing the estimate of U_{0k} in the first model and in the expanded model will inform whether these two characteristics explain the nurse measurement effects and whether they should be routinely collected in biosocial surveys.

R 3.5.2 was used for data cleaning and running the models. The models were estimated using the `rstanarm` package, an interface that facilitates the estimation of multilevel models using the Monte Carlo Markov Chain procedures from Stan. Estimation was implemented using four chains with 2000 iterations out of which the last 1000 were used for sampling. Weakly informative prior distributions were used: normal distribution for the intercept and slope and an exponential distribution for the residual.¹

Missing data was handled using listwise deletion. Over all five samples there was, on average, about 4.6% missing cases on the independent variables and 2.5% on the dependent variables. Over all the samples, an average of 8798 respondents, 4735 areas, and 116 nurses are used per survey. The full sample and descriptive statistics can be found in [Additional file 1.docx].

Results

Magnitude of nurse effects

To answer the first research question, a cross-classified model was estimated for each of the fourteen measurement outcomes of interest in each of the five waves of data collection. To facilitate the interpretation of the results, the nurse random effects are represented as Intraclass Correlation Coefficients (ICCs). These can be interpreted as the proportion of variation that is associated with each level of the data, in our case: nurses, areas, and residual variance after controlling for respondent characteristics.

Across all biomeasures collected in all surveys and waves, the average nurse effect on measurement is 13%, which is considered to be a medium-sized effect. This quantity varies by biomeasure with a minimum of 2% to

a maximum of 27%. Figure 1 shows how the nurse effects vary by type of measure collected. The largest nurse effects are for the physical performance measures: grip strength and lung capacity, while the smallest effects are for the anthropometric measures: pulse, height and weight. There also appears to be some variation within each measure depending on the order in which the measurements were collected. For example, the largest nurse effect for grip strength occurs at the first measurement, whereas the largest nurse effect for pulse is observed at the third measurement, although the differences are small.

Nurse effects on measurement error across surveys and over time

Next, we investigate if these nurse effects vary over time and data source. Figure 2 disentangles the variation by biomeasure, survey, and wave. Only small variations are observed: the average ICC in waves 2, 4, and 6 of ELSA are 13.9, 15.8, and 12.5%, respectively, whereas the average ICCs for the US wave 2 and BHPS wave 19 are 11.5 and 12.7%, respectively. For ELSA, it appears that the effects of nurses on lung capacity measures are higher in waves 2 to 4 compared to wave 6. One explanation for this difference is the change in the model of spirometer used in ELSA wave 6 [48]. On the other hand, waves 4 and 6 show higher levels of nurse effects on grip strength compared to wave 2, suggesting that nurse effects are worsening over time. For UKHLS the effects are more consistent although the average nurse effect is also higher for BHPS19 than for USW2.

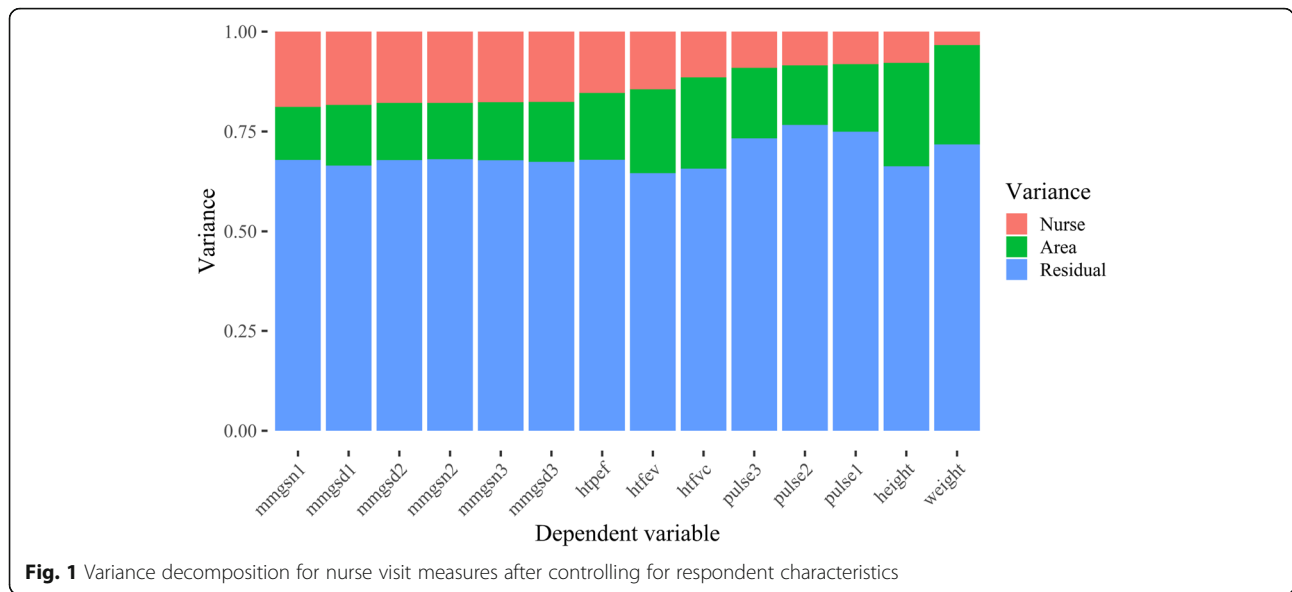
Comparing ELSAW2 with USW2 we see that nurse effects on lung capacity are higher in the former while the effect on grip strength is higher in the latter. On average, nurse effects are larger in ELSA (around 14%) than in the US and BHPS (around 12%).

Another way to visualize the differences in nurse effects over time and between data sources is shown in Fig. 3. Here, the higher levels of nurse effects for lung capacity in waves 2 and 4 of ELSA are more obvious. One can also see lower nurse effects on grip strength in ELSAW2 compared to the other data sources. Further, the figure confirms the low levels of nurse effects on the measures of pulse, height and weight in all data sources.

The impact of nurse characteristics on measurement error effects

Lastly, we investigate to what degree nurse characteristics (nurse experience and nurse age) explain the nurse measurement error effects found above. Overall, nurse characteristics explain only a very small proportion of variation. The average proportion of variation explained with nurse characteristics for each biomeasure ranges from a minimum of 0.1% to a maximum of 1.8%. Thus,

¹See <http://mc-stan.org/rstanarm/articles/priors.html> and <http://mc-stan.org/rstanarm/articles/glmr.html>

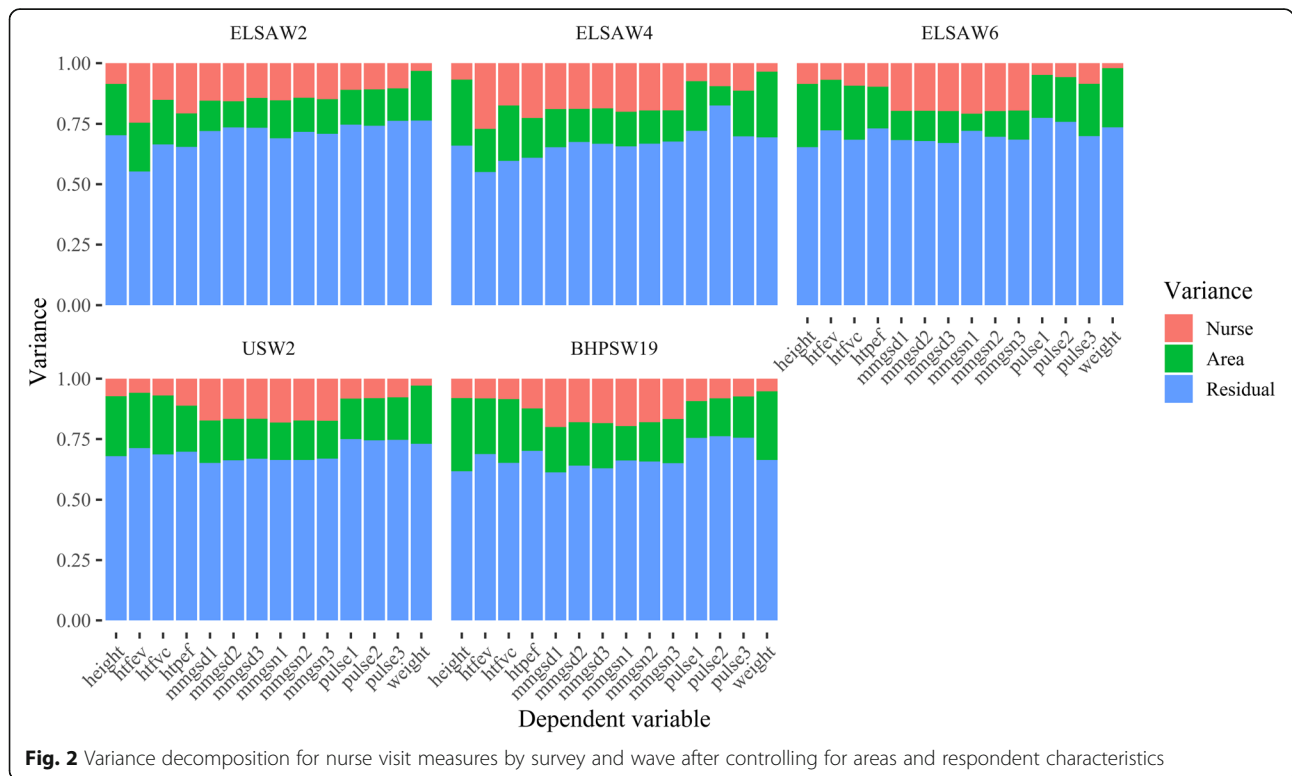


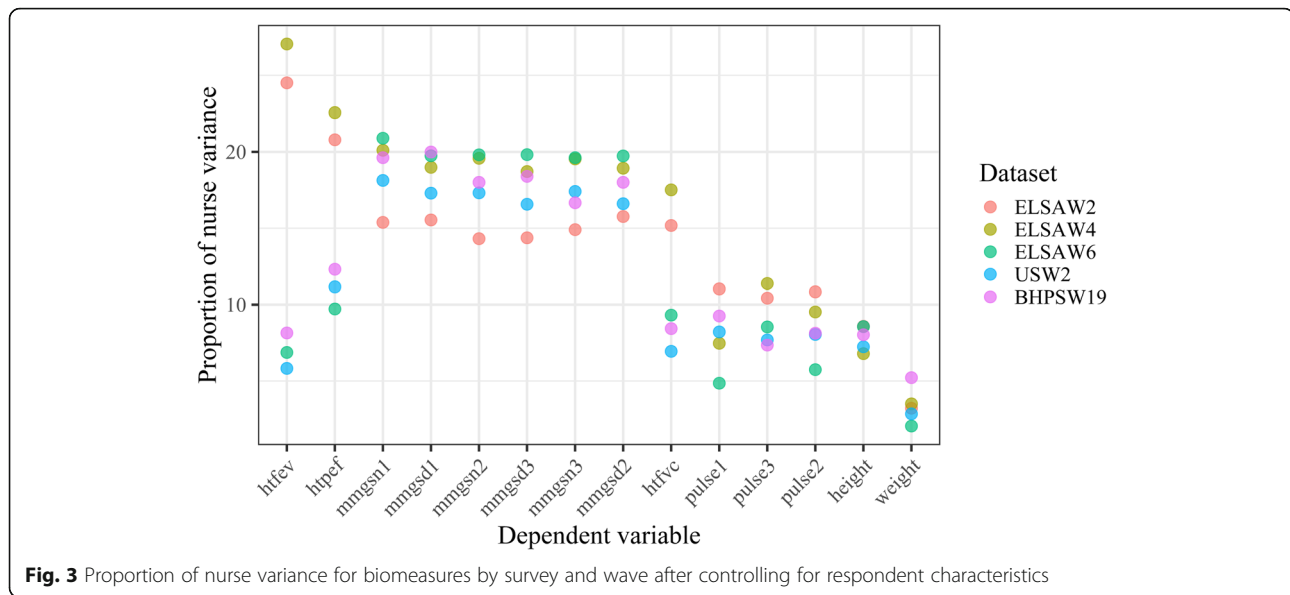
it appears that stronger background data are needed in order to understand the mechanisms that lead to nurse measurement effects and how to tackle them.

Discussion

This study aimed to investigate whether nurses influence the measurement of anthropometric measures as well as physical performance measures in large, nationally-

representative biosocial surveys. The investigation concentrated on 14 biomeasures collected in two surveys over five waves of data collection. Nurse effects were estimated using a cross-classified multilevel model that separated area and nurse effects and controlled for respondent characteristics. Overall, we found a medium-sized nurse effect. Across all biomeasures and data sources, nurses explained about 13% of the observed





variation, on average. This was shown to vary significantly from close to 0 % for some measures, such as pulse, height and weight and up to 27% for grip strength measurements. The finding that physical performance measures are more susceptible to nurse effects compared to anthropometric measures is consistent with other nurse-led studies, which have found measures of height and weight to be least prone to measurement error compared to other measures which require more careful administration and use of more sophisticated equipment [28].

Further, we investigated to what degree nurse effects vary by time and survey. This issue is important as varying nurse effects can distort estimates of change and comparisons between studies. Some differences over time and study were observed, but they did not reveal a consistent pattern. Nurse effects were larger for measures of lung capacity in ELSA compared to US and BHPS. On the other hand, nurse effects on grip strength were higher in US and BHPS compared to ELSA. Looking at comparisons over time it was found that ELSA wave 6 had smaller nurse effects on lung capacity compared to earlier waves 2 and 4, suggesting that nurse measurement improves over the course of the study for this particular measure. This pattern, however, was reversed for grip strength where the largest nurse effects were observed in the later waves of the study.

Finally, we attempted to explain these nurse effects using two important nurse characteristics measured in all five datasets: nurse experience and nurse age. Overall, the models explained only a very small proportion of the nurse effects (less than 2%) when the nurse characteristics were introduced. It is clear that additional nurse

characteristics are needed to better understand the processes that help explain measurement error effects.

This study has two important limitations. Firstly, in the absence of random allocation of nurses to respondents a modeling approach to control for confounding of nurse and area effects and respondent characteristics was used. Although this approach makes the assumption that both effects can be separated through control variables, it does yield good variation and highlights differences between biomeasures and data sources in nurse effects. Further, this is a standard approach used in investigating observer effects in observational studies [49]. Secondly, the analysis of nurse characteristics that explain nurse measurement effects is based on only two variables: experience and age. Although these variables are important from a theoretical viewpoint, we found that they have poor explanatory ability. Thus, more information about the nurses is needed to understand the processes underlying nurse measurement effects and allow researchers to control for these effects.

From these results, it is clear that making available the nurse ID variable as well as nurse characteristics can be useful not only for modelling nonresponse but also for estimating and accounting for nurse effects on measurement error. The results also suggest the need to perform sensitivity analyses that take into account the nurse effects on measurement when using data collected by nurses. These effects are similar to those found in the interviewer effects literature and highlight that objective measures of health collected by nurses are not impervious to measurement error.

Lastly, more research is needed to understand the mechanisms of nurse measurement error effects. This

can be done either through qualitative research or by collecting more detailed information about nurse attitudes and behaviors, similar to what has been done in the lay interviewer effects literature [19]. This can, in turn, inform procedures or training programs that can help minimize this unnecessary source of measurement error.

This research is among the first to investigate the effects of nurses on biological data collection in a population-representative household sample survey. This was informed by two distinct literatures: the medical research in biological data collection and the survey methodology research on interviewer effects. Our results indicate that this has the potential to be an important research area and opens up some intriguing research questions. For example, do these findings generalize to other types of health care professionals? Are these differences in nurse effects by type of measurement consistent across contexts and survey institutes? How do these effects influence substantive analyses? We hope this study will trigger avenues for research that can answer such questions.

Conclusions

Our study showed that some objective health measures collected by nurses in household surveys are susceptible to non-trivial amounts (up to 27%) of measurement error variability at the nurse level. Physical performance measures tended to be more influenced by nurse measurement error than anthropometric measures. Nurse characteristics did not explain a significant amount of the observed measurement error. We recommend that researchers account for this unnecessary source of error by incorporating nurse ID numbers in their analysis, or performing a sensitivity analysis to determine whether substantive findings are affected by nurse measurement effects. Moreover, a richer array of nurse characteristics should be included in biosocial data releases to further allow researchers to control for this.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12874-020-00922-2>.

Additional file 1. Descriptive statistics of dependent variables

Abbreviations

BHPS: British Household Panel Survey; BHPSW19: British Household Panel Survey, Wave 19; CM: Centimeters; ELSA: English Longitudinal Study of Ageing; ELSAW2: English Longitudinal Study of Ageing, Wave 2; ELSAW4: English Longitudinal Study of Ageing, Wave 4; ELSAW6: English Longitudinal Study of Ageing, Wave 6; HTFEV: Lung capacity amount of air that can be blown out in 1 sec; HTFVC: Lung capacity amount of air that can be forcibly blown out after a full inhalation; HTPEV: Lung capacity speed of air moving out of lungs; ICC: Intraclass Correlation Coefficient; KG: Kilograms; MMGSD: Grip strength for dominant hand; MMGSN: Grip strength for non-dominant hand; UKHLS: UK Household Longitudinal Study;

US: Understanding Society – UK Household Longitudinal Study; USW2: Understanding Society – UK Household Longitudinal Study, Wave 2

Acknowledgements

This paper benefited from feedback received during a presentation given at the Munich Center for the Economics of Aging (MEA) at the Max-Planck-Institute for Social Law and Social Policy, 13 February 2019. This paper was supported by a grant from the National Centre for Research Methods (NCRM) of the Economic and Social Research Council: [ES/L008351/10].

Authors' contributions

Both authors contributed to the research questions and study design. AC carried out the analysis and drafted the manuscript. JWS drafted the manuscript and provided critical revisions. The author(s) read and approved the final manuscript.

Funding

Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex. The English Longitudinal Study of Ageing is supported by the National Institute on Aging (NIA-NIH) USA (grant number 5 R01 AG017644–16) and a consortium of the United Kingdom government departments coordinated by the Office for National Statistics. The work of AC was supported by a grant from the National Centre for Research Methods (NCRM) of the Economic and Social Research Council: [ES/L008351/1]. None of these funding bodies played a role in the design of the present study and collection, analysis, and interpretation of data and in the writing of the manuscript.

Availability of data and materials

Data from Understanding Society (US) and the English Longitudinal Study of Ageing (ELSA) are available from the UK Data Service for researchers who meet the criteria for access to confidential data, under conditions of the End User License <http://ukdataservice.ac.uk/media/455131/cd137-enduserlicence.pdf>. The Understanding Society data can be accessed from: <https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000053>. The English Longitudinal Study of Ageing data can be accessed from: <http://discover.ukdataservice.ac.uk/series/?sn=200011>. Contact with the UK data service regarding access to Understanding Society and the English Longitudinal Study of Ageing can be made through the website <http://ukdataservice.ac.uk/help/get-in-touch.aspx>, by phone + 44 (0)1206 872143, or by email at help@ukdataservice.ac.uk.

Ethics approval and consent to participate

The University of Essex Ethics Committee has approved all data collection on Understanding Society main study and innovation panel waves. Approval for the collection of biosocial data by trained nurses was obtained from the National Research Ethics Service (Understanding Society - UK Household Longitudinal Study: A Biosocial Component, Oxfordshire A REC, Reference: 10/H0604/2). Ethical approval for all the ELSA waves was granted from NHS Research Ethics Committees under the National Research and Ethics Service (NRES). Informed consent was obtained from all participants in both studies.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests

Author details

¹University of Manchester, Humanities Bridgford Street, Manchester M13 9PL, UK. ²Ludwig Maximilian University of Munich, Munich, Germany. ³University of Mannheim, Mannheim, Germany. ⁴Institute for Employment Research, Regensburger Strasse 104, 90478 Nuremberg, Germany.

Received: 7 July 2019 Accepted: 12 February 2020

References

- Sakshaug JW, Ofstedal MB, Guyer H, Beebe T. The collection of biospecimens in health surveys. In: T. Johnson (Ed.). *Handbook of Health Survey Methods*. Hoboken: Wiley; pp. 383–419.
- Lindau ST, McDade TW. Minimally invasive and innovative methods for biomeasure collection in population-based research. US: National Academies Press; 2008. [cited 2019 Mar 25]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK62423/>.
- McFall SL, Booker C, Burton J, Conolly A. Implementing the biosocial component of understanding society–nurse collection of biomeasures. University of Essex: Institute for Social and Economic Research; 2012.
- Sakshaug JW, Couper MP, Ofstedal MB. Characteristics of physical measurement consent in a population-based survey of older adults. *Med Care*. 2010 Jan;48(1):64–71.
- Schonlau M, Reuter M, Schupp J, Montag C, Weber B, Dohmen T, et al. Collecting genetic samples in population wide (Panel) surveys: feasibility, nonresponse and selectivity. *Surv Res Methods*. 2010;4(2):121–6 [cited 2019 Mar 25] Available from: <https://ojs.ub.uni-konstanz.de/srm/article/view/3959>.
- National Research Council. Cells and surveys: should biological measures be included in social science research? . 2000 [cited 2019 Mar 25]. Available from: <https://www.nap.edu/catalog/9995/cells-and-surveys-should-biological-measures-be-included-in-social>.
- National Research Council. Biosocial surveys. 2007 [cited 2019 Mar 25]. Available from: <https://www.nap.edu/catalog/11939/biosocial-surveys>.
- National Research Council. Conducting biosocial surveys: collecting, storing, accessing, and protecting biospecimens and biodata . 2010 [cited 2019 Mar 25]. Available from: <https://www.nap.edu/catalog/12942/conducting-biosocial-surveys-collecting-storing-accessing-and-protecting-biospecimens-and>.
- Whelan BJ, Savva GM. Design and methodology of the Irish longitudinal study on ageing. *J Am Geriatr Soc*. 2013 May;61(Suppl 2):S265–8.
- Zipf G, Chiappa M, Porter KS, Ostchega Y, Lewis BG, Dostal J. National health and nutrition examination survey: plan and operations, 1999–2010. *Vital Health Stat Ser 1 Programs Collect Proc*. 2013;56:1–37.
- Kearney PM, Cronin H, O'Regan C, Kamiya Y, Savva GM, Whelan B, et al. Cohort profile: the Irish longitudinal study on ageing. *Int J Epidemiol*. 2011 Aug;40(4):877–84.
- Fisher GG, Ryan LH. Overview of the health and retirement study and introduction to the special issue. *Work Aging Retire*. 2018 Jan;4(1):1–9.
- Harris KM, Halpern CT, Whitsel E, Hussey J, Tabor J, Entzel P, et al. The national longitudinal study of adolescent to adult health: research design [Internet]. Available from: <http://www.cpc.unc.edu/projects/addhealth/design>.
- Herd P, Carr D, Roan C. Cohort Profile: Wisconsin longitudinal study (WLS). *Int J Epidemiol*. 2014;43(1):34–41 [cited 2019 Mar 25] Available from: <https://academic.oup.com/ije/article/43/1/34/729687>.
- Jaszczak A, Lundeen K, Smith S. Using nonmedically trained interviewers to collect biomeasures in a national in-home survey. *Field Methods*. 2008;21(1):26–48. [cited 2019 Mar 25] Available from: <https://doi.org/10.1177/1525822X08323988>.
- Malter F, Börsch-Supan A. SHARE wave 6: panel innovations and collecting dried blood spots. Munich: MEA, Max Planck Institute for Social Law and Social Policy; 2017.
- Blake M, Marmot M, Phelps A, Steptoe A, Oldfield Z, Clemens S, et al. English Longitudinal Study of Ageing: Waves 0–8, 1998–2017 [Internet]. UK Data Service; 2018 [cited 2018 Oct 17]. Available from: <https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=5050#15>.
- Guyer H, Ofstedal MB, Lessof C, Cox K. The benefits and challenges of collecting physical measures and biomarkers in cross-national studies. *Ann Arbor, MI: Institute for Social Research*; 2017. Available from: https://hrs.isr.umich.edu/sites/default/files/biblio/Collecting%20PM-Bio%20Data_DocumentationReport.pdf.
- Korbmacher JM. Interviewer effects on respondents' willingness to provide blood samples in SHARE. SHARE working paper series 20–2014. Munich: MEA, Max Planck Institute for Social Law and Social Policy; 2014.
- SL MF, Conolly A, Burton J. Collecting Biomarkers Using Trained Interviewers. Lessons Learned from a Pilot Study. *Surv Res Methods*. 2014; 8(1):57–66 [cited 2019 Mar 25] Available from: <https://ojs.ub.uni-konstanz.de/srm/article/view/5471>.
- Ali S, Rouse A. Practice audits: reliability of sphygmomanometers and blood pressure recording bias. *J Hum Hypertens*. 2002;16(5):359–61.
- Andersen UO, Henriksen JH, Jensen G. Copenhagen City heart study group. Sources of measurement variation in blood pressure in large-scale epidemiological surveys with follow-up. *Blood Press*. 2002;11(6):357–65.
- Bogan B, Kritzer S, Deane D. Nursing student compliance to standards for blood pressure measurement. *J Nurs Educ*. 1993;32(2):90–2.
- Dickson BK, Hajjar I. Blood pressure measurement education and evaluation program improves measurement accuracy in community-based nurses: a pilot study. *J Am Acad Nurse Pract*. 2007;19(2):93–102.
- Grim CM, Grim CE. A curriculum for the training and certification of blood pressure measurement for health care providers. *Can J Cardiol*. 1995; 11(Suppl H):38H–42H.
- Selenta C, Hogan BE, Linden W. How often do office blood pressure measurements fail to identify true hypertension? An exploration of white-coat normotension. *Arch Fam Med*. 2000 Jun;9(6):533–40.
- Trilling JS, Froom J. The urgent need to improve hypertension care. *Arch Fam Med*. 2000;9(9):794–801 [cited 2019 Mar 25] Available from: <http://europepmc.org/abstract/med/11031384>.
- Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr*. 1999 Sep;82(3):165–77.
- Ulijaszek SJ, Lourie JA. Intra- and inter-observer error in anthropometric measurement. In: Ulijaszek SJ, CGN M-T, editors. *Anthropometry: the Individual and the Population*. Cambridge: Cambridge University Press; 2005. p. 30–55.
- Armstrong RS. Nurses' knowledge of error in blood pressure measurement technique. *Int J Nurs Pract*. 2002;8(3):118–26.
- Bague J-P, de Poli F, Sosner P, Jaboureck O, Nait-Saidi L, Vautrin E, et al. Using a large cuff reduces the difference between peripheral and central blood pressure readings. The BP-CUFF study. *Int J Cardiol*. 2013;170(2):e43–4 [cited 2019 Mar 25] Available from: [https://www.internationaljournalofcardiology.com/article/S0167-5273\(13\)01916-5/abstract](https://www.internationaljournalofcardiology.com/article/S0167-5273(13)01916-5/abstract).
- Baillie L, Curzio J. A survey of first year student nurses' experiences of learning blood pressure measurement. *Nurse Educ Pract*. 2009;9(1):61–71.
- Bur A, Herkner H, Vlcek M, Woisetschläger C, Derhaschnig U, Delle Karth G, et al. Factors influencing the accuracy of oscillometric blood pressure measurement in critically ill patients. *Crit Care Med*. 2003;31(3):793–9.
- Kay LE. Accuracy of blood pressure measurement in the family practice center. *J Am Board Fam Pract*. 1998 Aug;11(4):252–8.
- Maxwell MH, Waks AU, Schroth PC, Karam M, Dornfeld LP. Error in blood-pressure measurement due to incorrect cuff size in obese patients. *Lancet Lond Engl*. 1982;2(8288):33–6.
- Canner PL, Borhani NO, Oberman A, Cutler J, Prineas RJ, Langford H, et al. The hypertension prevention trial: assessment of the quality of blood pressure measurements. *Am J Epidemiol*. 1991 Aug 15;134(4):379–92.
- Gillespie A, Curzio J. Blood pressure measurement: assessing staff knowledge. *Nurs Stand R Coll Nurs G B*. 1987. 1998;12(23):35–7 [cited 2019 Mar 25] Available from: <http://europepmc.org/abstract/med/9528543>.
- Torrance C, Serginson E. An observational study of student nurses' measurement of arterial blood pressure by sphygmomanometry and auscultation. *Nurse Educ Today*. 1996;16(4):282–6.
- Pashazadeh F, Cernat A, Sakshaug JW. The effects of biological data collection in longitudinal surveys on subsequent wave cooperation. In: *methodology of longitudinal surveys volume II*. Chichester: Wiley; in press. University Of Essex IFS. Understanding Society: Waves 1–7, 2009–2016 and Harmonised BHPS: Waves 1–18, 1991–2009: UK Data Service; 2018. [cited 2018 Oct 17] Available from: <https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=6614#11>.
- Knies G. Understanding society. UK Househ Longitud Study Waves; 2018. p. 1–7.
- McFall S, Petersen J, Kaminska O, Lynn P. Waves 2 and 3 Nurse Health Assessment, 2010–2012. Guide to Nurse Health Assessment [Internet]. Colchester, Essex: Institute for Social and Economic Research; (Understanding Society The UK Household Longitudinal Study); 2014. p. 1–38. Available from: https://www.understandingsociety.ac.uk/d/100/7251_User_Guide_Health_Assmt_w2_w3.pdf?1392855567.
- Bridges S, Hussey D, Blake M. The dynamics of ageing: The 2012 English Longitudinal Study of Ageing (Wave 6): NatCen; 2015. p. 1–100.
- Cheshire H, Hussey D, Medina J, Piekerling K, Wood N, Ward K, et al. Financial circumstances, health and well-being of the older population in England: The 2008 English Longitudinal Study of Ageing; 2012. p. 1–76.

45. Scholes S, Taylor R, Cheshire H, Cox K, Lessof C. Retirement, health and relationships of the older population in England: The 2004 English Longitudinal Study of Ageing; 2008. p. 1–101.
46. Taylor R, Conway L, Calderwood L, Lessof C, Cheshire H, Cox K, et al. Health, wealth and lifestyles of the older population in England: The 2002 English Longitudinal Study of Ageing; 2007. p. 1–74.
47. NatCen Social Research. English Longitudinal Study of Ageing (ELSA) Wave 2, 4 and 6 User Guide to the nurse datasets. London, England: NatCen Social Research; 2015. p. 1–28.
48. NatCen Social Research. User Guide to the Nurse Visit Datasets. Waves 2, 4, 6, 8. London, England: NatCen Social Research; 2018. p. 1–24.
49. West BT, Blom AG. Explaining interviewer effects: a research synthesis. *J Surv Stat Methodol*. 2017;5(2):175–211.
50. Snijders TAB, Bosker R. *Multilevel Analysis: An Introduction To Basic And Advanced Multilevel Modeling*. 2nd ed: Sage Publications Ltd; 2011.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

