How to….obtain accurate objective measurements of health at a distance

Brett, C. E. & Deary, I. J.

Abstract

Large scale longitudinal studies are an excellent tool for increasing our understanding of the aetiology of health and disease. Obtaining accurate measures of health status is important in these kinds of studies. However, self-report measures of health are subject to bias and obtaining objective health measures can be costly. This paper outlines the process and challenges of designing a home testing kit to enable participants to obtain objective health measures themselves, using the example of a new cohort study, the 6-Day Sample.

Introduction: the problem

Longitudinal cohort or population-based studies are an invaluable tool for addressing epidemiological questions around the causes and correlates of health (Cooper et al, 2012, Kuh et al, 2003, Pearson, 2011). Asking a large number of individuals the same questions increases statistical power and enables researchers to investigate the influence of confounding variables such as age, sex, geography, educational and occupational level. Obtaining a measure of health status – both past and present – is essential.

Traditionally, large scale studies have favoured self-report as a means of collecting this data, with a considerable literature on the reliability and predictive validity of self-reported health (Idler & Benyamini, 1997, DeSalvo et al, 2006, Kuhn et al, 2006). However, obtaining accurate self-report measures of current health status can be problematic. It is well known that people tend to over-estimate participation in positive health behaviours such as physical activity (Prince et al, 2008) or healthy eating (Schoeller, 1990) and underestimate negative health behaviours such as drinking (Stockwell et al, 2004) and smoking (Connor-Gorber et al, 2009). Even simple physical measures such as height and weight are often reported inaccurately. Height is generally over-estimated,

particularly among shorter (Stewart, 1982, Gunnell et al, 2000) and overweight adults (Rowland, 1990, Vailas & Nitzke, 1998). Older adults frequently over-estimate their height, perhaps reporting height from an earlier age (Gunnell et al, 2000, Sahyoun et al, 2008, Shiely et al, 2013). Weight is often inaccurately reported (Rowland, 1990, Vailas & Nitzke, 1998, Sahyoun et al, 2008, Shiely et al, 2013). As Body Mass Index (BMI), which is calculated from height and weight, is a vital indicator of health status, these self-report inaccuracies can have implications for a study's findings.

To overcome these limitations, researchers have placed increasing emphasis on objective measurements of health, including so-called biomarkers, which embody sub-clinical changes in health which might act as precursors to disease and poor health. In aging research, biomarkers represent the underlying biological and physiological processes of aging, enabling researchers to calculate a more accurate measure of an individual's physiological and functional age than chronological age alone (Baker & Sprott, 1988, Anstey & Smith, 1999, Sprott, 2010). In order to collect objective health measurements, many studies, such as the English and Scottish Health Surveys[1] and the UK Biobank, (Allen et al, 2012) now conduct extensive physical and/or cognitive testing of their participants. This approach maximises the quality of the data collected through detailed protocols and careful training, and allows for the safe collection of blood and other biological samples.

But for many studies this approach is impossible due to costs or practicalities, or inappropriate due to the nature of the sample population. Researchers are looking to find new ways of obtaining accurate, objective measures of physical health that can be administered by individuals themselves with minimal expenditure.

Methods: the 6-Day Sample study

[1] http://healthsurvey.hscic.gov.uk/support-guidance/public-health/health-survey-for-england.aspx, http://www.scotland.gov.uk/Topics/Statistics/Browse/Health/scottish-health-survey

The current article outlines the measures used in a new longitudinal cohort study, the 6-Day Sample (Deary, Whalley, & Starr, 2009, Brett & Deary, 2014). This group of 1208 individuals, born on 6 days of 1936, were chosen as a representative sample of the Scottish population. They were followed up from age 11 to 27, providing comprehensive information on their intelligence, family, personal characteristics and educational and occupational attainment. This group has been traced and survivors living throughout Scotland, England, and Wales were invited to a follow-up study of health and wellbeing in old age. The study investigators were keen to obtain multiple markers of physical health, which, due to the geographical spread of participants, their age and the potential burden of travelling to Edinburgh for testing, needed to be conducted by participants themselves at home.

Choosing what (and what not) to measure and devising a testing kit.

The decision as to which measures to include should be informed by the study's research questions and hypotheses. These choices are often made by principal investigators well in advance of applying for research funding or ethical approval. Implementing their suggestions and selecting a suitable set of tests for inclusion in the study protocol can require tact and diplomacy. Consideration needs to be given to practicalities, with only the most accurate, effective and least burdensome tests chosen. There is little point devising a complex and ingenious testing kit if participants will take one look and throw it away.

The 6-Day Sample study consists of eight workstreams each with unique research questions and hypotheses. These included genetics, stress reactivity, and general health. The first two required the collection of saliva samples for DNA and cortisol analysis (Heaney et al, 2010). A range of biomarkers were chosen to measure general health, all of which have been related to morbidity and mortality in older adults: BMI (Rantanen et al, 2000), waist-hip ratio (Price et al, 2006, Srikanthan et al, 2009), leg length (Gunnell et al, 1998), balance (Berkman et al, 1993, Klein et al, 2005)

function (Cook et al, 1993, Cook et al, 1995, Knudtson et al, 2009), and arm strength (Doherty, 2003, Manor et al, 2006).

Sourcing instructions and equipment

It is essential that all participants complete the tests in the same way. This ensures the validity of the data and protects the study's scientific integrity. The physical testing 'kit' should be designed to maximise the consistency and reliability of results across participants of varying levels of functional and cognitive ability. As participants will complete these tests without the presence of a researcher, it is essential that the tests are easy to use, with clear and accurate instructions, and that participants are provided with everything they need to complete the tests correctly.

The methods sections of relevant papers can be a useful source of instructions, and equipment manufacturers will often be willing to provide sample equipment alongside accurate instructions. The internet can be a useful source if used with caution and a well-thought-out search strategy. Ideally, all instructions used should be replicated across multiple sources.

When devising the kit, it is good practice to be prepared for every eventuality. Simple things can make the difference between success and failure. For example, ensuring correct postage and addresses are on all return envelopes, packaging the kit carefully to minimise damage in transit, providing suitable packaging for the safe (and legal) return of biological samples, ensuring study contact details are clearly visible in case participants have questions, designing a user-friendly answer sheet that allows flexibility while minimising error, and making sure that all items returned from each participant can be identified.

Trial all equipment and instructions – on yourself, colleagues, and age-appropriate individuals – before finalising the testing kit and committing to bulk purchases. In the 6-Day Sample study, peak

flow meters from two manufacturers were purchased and their reliability compared. The whole testing kit was trialled by willing age-appropriate volunteers, who were encouraged to give honest and detailed feedback.

The contents of the 6-Day Sample testing kit are shown in Table 1.

Measure	**Test**	**Equipment provided**
BMI	Height / Weight	Pencil 3m flexible tape measure
Body morphometry	Waist, hip & calf circumference	Tape measure
	Leg length	
Arm strength	Elbow flexions	1.5m low resistance TheraBand Countdown timer
Balance	One leg stand	Countdown timer
Lung function	Peak expiratory flow	Peak flow meter
Genetics	Saliva sample	Oragene testing kit
Stress reactivity	Salivary cortisol	3 x Salivette cortisol tubes
General	N/A	Branded pen & badge, 3 x postage paid return envelopes, 12-page A5 instruction booklet, DVD

Table 1: Contents of the 6-Day Sample physical testing kit

Creating an instructional video

One way of facilitating the correct completion of physical tests is to demonstrate them in a video. This should not be undertaken lightly. The costs of producing a video professionally may be beyond the reach of many research studies, and completing the task in-house – as was done for the 6-Day Sample study – is time-consuming. However, the benefits in terms of participation and consistency are plentiful.

The video should cover each and every aspect of the physical testing measurements, preferably demonstrated by someone close to the participants' age. Creating a storyboard and script in advance facilitates this process and ensures that sufficient and appropriate footage is shot. The script should be economical and match the written instructions. The footage itself should be shot against a plain background to reduce distraction, using the same testing kit that participants will receive. If editing the video yourself, allow plenty of time – especially if inexperienced. The 6-Day Sample video took 2 weeks to edit 30 minutes of footage. Still shots or pictures are useful in instances when the narrative outruns the video footage.

Quality control

Participants may still make mistakes while completing the physical testing. Documented quality control measures are needed to check the integrity of the data received. If samples are being collected, these will often be processed at a laboratory with their own, robust, quality control checks in place. For other measures, a simple check involves looking for, and possibly removing, impossible values or extreme outliers.

Results

The 6-Day Sample study achieved a return rate of 27.3%. Comparison of childhood intelligence data revealed a one standard deviation difference in intelligence between responders and refusals (IQ = 115.59 vs 99.87, $F=15.037$, $p<.001$). Most participants completed all the physical measures, with very few impossible values, and the vast majority of genetic and cortisol samples received were of sufficient quality for analysis. Ten participants were also members of the Lothian Birth Cohort (LBC) 1936 (Deary et al, 2007), and had undergone extensive physical and cognitive testing within a year of our study. Their results from the equivalent measures within the two studies were compared (Table 2).

Measure	LBC Wave 3	6-Day Sample	Correlation	p
Height (cm)	165.95	169.44	.841	.002
Weight (kg)	78.00	78.85	.761	.011
Peak expiratory flow	320.83	381.93	.703	.023
Arm strength*			.389	.266

Table 2: comparison of LBC and 6-Day Sample measures. N = 10.

*LBC = grip strength (dominant hand), 6-Day Sample = number of elbow flexions completed in 30 seconds.

Consistent with previous research, height was overestimated and weight underestimated.

Conclusion and implications

Collecting objective measures of current health status is important for large-scale studies. The experience of the 6-Day Sample study has shown that valid measures can be obtained at a distance. Despite our best efforts, the observed difference in intelligence between participants and non-participants suggests that some potential participants may have been put off by the cognitive complexity of completing these tests (and a lengthy questionnaire) at home (Bowling, 2005). Additionally, the discrepancies noted in a small group between the self-administered and clinic-administered measures suggest that further validation of the physical tests may be required.

Acknowledgements

The 6-Day Sample study and CEB are supported by a UK cross council Lifelong Health and Wellbeing Initiative grant (MRCG1001401). Both authors are members of the Centre for Cognitive Ageing and Cognitive Epidemiology at the University of Edinburgh, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged.

References

Allen, N., Sudlow, C., Downey, P., Peakman, T., Danesh, J., Elliott, P., Gallacher, J., Green, J., Matthews, P., Pell, J., Sprosen, T., & Collins, R. (2012). UK Biobank: Current status and what it means for epidemiology. *Health Policy and Technology, 1(3),* 123-126.

Anstey, K. J., & Smith, G. A. (1999). Interrelationships among biological markers of aging, health, activity, acculturation, and cognitive performance in late adulthood. *Psychology of Aging, 14,* 605-618.

Baker, G. T., & Sprott, R. L. (1988). Biomarkers of aging. *Experimental Gerontology, 23,* 223-239.

Berkman, L. F., Seeman, T. E., Albert, M., Blazer, D., Kahn, R., Mohs, R., Finch, C., Schneider, E., Cotman, C., McClearn, G., Nesselroade, J., Featherman, D., Garmezy, N., McKhann, G., Brim, G., Prager, D., & Rowe, J. (1993). High, usual and impaired functioning in community-dwelling older men and women: Findings from the MacArthur Foundation research network on successful aging. *Journal of Clinical Epidemiology, 46(10),* 1129-1140.

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health, 27(3),* 281-291.

Brett, C. E., & Deary, I. J. (2014). The long road to realising data linkage: obtaining permission for the 6-Day Sample study. *Longitudinal and Life Course Studies, 5,* 283-298.

Connor-Gorber, S., Schofield-Hurwitz, S., Hardt, J., Levasseur, G., & Tremblay, M. (2009). The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. *Nicotine & Tobacco Research, 11(1),* 12-24.

Cook, N. R., Albert, M. S., Berkman, L. F., Blazer, D., Taylor, J. O., & Hennekens, C. H. (1995). Interrelationships of peak expiratory flow rate with physical and cognitive function in the elderly: MacArthur Foundation studies of aging. *Journal of Gerontology, 50A(6),* M317-M323.

Cook, N. R., Evans, D. A., Scherr, P. A., Speizer, F. E., Taylor, J. O., & Hennekens, C. H. (1993). Peak expiratory flow rate and 5-year mortality in an elderly population. *American Journal of Epidemiology, 133(8),* 784-794.

Cooper, R., Frank, J., Leyland, A., Hardy, R., Lawlor, D. A., Wareham, N. J., Dezateux, C., & Inskip, H. (2012). Using cohort studies in lifecourse epidemiology. *Public Health, 126*, 190-192.

Deary, I. J., Gow, A. J., Taylor, M. D., Corley, J., Brett, C., Wilson, V., Campbell, H., Whalley, L. J., Visscher, P. M., Porteous, D. J., & Starr, J. M. (2007). The Lothian Birth Cohort 1936: a study to examine the influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatrics, 7*, 28.

Deary, I. J., Whalley, L. J., & Starr, J. M. (2009). *A lifetime of intelligence: follow-up studies of the Scottish Mental Surveys of 1932 and 1947*. Washington, DC: American Psychological Association.

DeSalvo, K.B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2006). Mortality prediction with a single general self-rated health question. *Journal of General Internal Medicine, 21(3),* 267-275.

Doherty, T. J. (2003). Invited review: Aging and sarcopenia. *Journal of Applied Physiology, 95,* 1717-1727.

Gunnell, D. J., Berney, L., Holland, P., Maynard, M., Blane, D., Frankel, S., & Davey Smith, G. (2000). How accurately are height, weight and leg length reported by the elderly, and how closely are they related to measurements recorded in childhood? I*nternational Journal of Epidemiology, 29*, 456-464.

Gunnell, D. J., Davey Smith, G., Holly, J. M. P., & Frankel, S. (1998). Leg length and risk of cancer in the Boyd Orr cohort. *BMJ, 317*, 1350-1351.

Heaney, J. L. J., Phillips, A. C., & Carroll, D. (2010). Ageing, depression, anxiety, social support and the diurnal rhythm and awakening response of salivary cortisol. *International Journal of Psychophysiology, 78(3),* 201-208.

Hedden. T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nature Reviews Neuroscience, 5,* 87–96.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior, 38(1*), 21-37.

Klein, B. E. K., Klein, R., Knudtson, M. D., & Lee, K. E. (2005). Frailty, morbidity and survival. *Archives of Gerontology and Geriatrics, 41*, 141-149.

Knudtson, M. D., Klein, B. E. K., & Klein, R. (2009). Biomarkers of aging and falling: the Beaver Dam eye study. *Archives of Gerontology and Geriatrics, 49(1),* 22-26.

Kuh, D., Ben-Shlomo, Y., Lynch, J., Hallqvist, J., & Power, C. (2003). Life course epidemiology. *Journal of Epidemiology and Community Health, 57*, 778-783.

Kuhn, R., Rahman, O., & Menken, J. (2006). Survey measures of health: How well do self-reported and observed indicators measure health and predict mortality? B. Cohen, & J. Menken (eds). *Aging in Sub-Saharan Africa: Recommendation for Furthering Research*. Washington (DC): National Academics Press.

Manor, B., Topp, R., & Page, P. (2006). Validity and reliability of measurements of elbow flexion strength obtained from older adults using elastic bands. *Journal of Geriatric Physical Therapy, 29(1),* 6.

Pearson, H. (2011). Study of a lifetime. *Nature, 471*, 20-24.

Price, G. M., Uauy, R., Breeze, E., Bulpitt, C. J., & Fletcher, A. E. (2006). Weight, shape and mortality risk in older persons: elevated waist-hip ratio, not high body mass index, is associated with a greater risk of death. *American Journal of Clinical Nutrition, 84,* 449-460.

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Connor Gorber, S., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a

systematic review. *International Journal of Behavioral Nutrition and Physical Activity, 5*, 56.

Rantanen, T., Harris, T., Leveille, S. G., Visser, M., Foley, D., Masaki, K., & Guralnik, J. M. (2000). Muscle strength and body mass index as long-term predictors of mortality in initially healthy men. *Journal of Gerontology, 55A(3),* M168-M173.

Rowland, M. L. (1990). Self-reported weight and height. *American Journal of Clinical Nutrition, 52*, 1125-1133.

Sahyoun, N. R., Maynard, L. M., Zhang, X. L., & Serdula, M. K. (2008). Factors associated with errors in self-reported height and weight in older adults. *The Journal of Nutrition, Health & Aging, 12(2),* 108-115.

Schoeller, D. A. (1990). How accurate is self-reported dietary energy intake? *Nutrition Review, 48(10*), 373-379.

Shiely, F., Hayes, K., Perry, I. J., & Kelleher, C. C. (2013). Height and weight bias: The influence of time. *PLOS ONE, 8(1),* e54386.

Srikanthan, P., Seeman, T. E., & Karlamangla, A. S. (2009). Waist-hip-ratio as a predictor of all-cause mortality in high-functioning older adults. *Annals of Epidemiology, 19(10),* 724-731.

Sprott, R. L. (2010). Biomarkers of aging and disease: Introduction and definitions. *Experimental Gerontology, 45,* 2-4.

Stewart, A. L. (1982). The reliability and validity of self-reported weight and height. *Journal of Chronic Disease, 35*, 295-309.

Stockwell, T., Donath, S., Cooper-Stanbury, M., Chikritzhs, T., Catalano, P., & Mateo, C. (2004). Under-reporting of alcohol consumption in household surveys: a comparison of quantity-frequency, graduated-frequency and recent recall. *Addiction, 99(8),* 1024-1033.

Vailas, L. I., & Nitzke, S. A. (1998). Self-reported versus measured weight and height in an older adult meal program population. *Journal of Gerontology, 53A(6),* M481-M483.