CONTENTS

78 EDITORIAL – Deborah Loxton

80 Getting started: ‘Preparing the ground’ and ‘planting the vines’ for longitudinal research – Julie Ellen Byles, Annette Dobson, Lois Bryson and Wendy J Brown

92 Human resources for longitudinal studies: Matching people to skills and tasks – Penny Warner-Smith, Deborah Loxton and Wendy J Brown

104 Accessing and disseminating longitudinal data: Protocols and policies – Catherine Chojenta, Rosemary Mooney and Penny Warner-Smith

114 Longitudinal survey development and design – Deborah Loxton and Anne Young

126 Recruiting for a longitudinal study: Who to choose, how to choose and how to enhance participation – Lyn Adamson, Anne Young and Julie Ellen Byles

137 Developing relationships and retaining participants in a longitudinal study – Lyn Adamson and Catherine Chojenta

147 Cohort management: Developing and maintaining participant databases in longitudinal studies – Lyn Adamson and Anna Graves

156 Data management: The building blocks of clean and accurate longitudinal datasets – Anna Graves, Jean Ball and Eliza Fraser

175 Working with longitudinal data: Attrition and retention, data quality, measures of change and other analytical issues – Anne Young, Jennifer Powers and Virginia Wheway

187 Conducting substudies in a longitudinal research project – Jenny Helman, Deborah Loxton, Lyn Adamson, Anna Graves and Jennifer Powers

199 Communication and dissemination of longitudinal study findings – Catherine Chojenta, Julie Ellen Byles, Deborah Loxton and Rosemary Mooney
This special issue of the *International Journal of Multiple Research Approaches* is a peer-reviewed practical guide to the development and management of longitudinal studies written by researchers and personnel who have been running the Australian Longitudinal Study on Women’s Health for over ten years.

Internationally, longitudinal research has become increasingly important to both the academic community and state policy makers. Longitudinal studies allow causal relationships to be examined and are critical to understanding issues associated with ageing. While the famous ‘Seven Up’ TV series is probably the best known, longitudinal studies can include many thousands of participants (eg the UK ‘Million Women’ study). However, there is currently a gap in the research methodology literature for a single volume that offers direction and advice about how the practical challenges of running longitudinal research studies can be met. This series endeavours to address this gap in knowledge using the Australian Longitudinal Study on Women’s Health as a case study throughout the eleven papers.

The Australian Longitudinal Study on Women’s Health (also known as Women’s Health Australia) is a population-based mailed survey, which examines the health of Australian women over a 20 year period. The project was designed to explore factors that influence health among women who are broadly representative of the entire Australian population and takes a comprehensive view of all aspects of health throughout the lifespan. The first main data collection occurred in 1996, when a random sample of over 40,000 women were recruited into three cohorts, then aged 18–23, 45–50 and 70–75 years. Subsequent data collections were

| Table 1: Survey data collection schedule used by the Australian Longitudinal Study on Women’s Health |
|--------|--------|--------|--------|--------|--------|--------|--------|
|       | Survey 1 | Survey 2 | Survey 3 | Survey 4 | Survey 5 | Survey 6 | Survey 7 |
|         | 18-23 yrs| 22-27 yrs| 25-30 yrs| 28-33 yrs| 31-36 yrs| 34-39 yrs| 37-42 yrs|
|         | 45-50 yrs| 47-52 yrs| 50-55 yrs| 53-58 yrs| 56-61 yrs| 59-64 yrs| 62-67 yrs|
|         | 70-75 yrs| 73-78 yrs| 76-81 yrs| 79-84 yrs| 82-87 yrs| 85-90 yrs| 88-93 yrs|
undertaken so that one cohort is now surveyed each year, according to the schedule provided in Table 1.

While longitudinal studies offer many benefits, including the ability to assess the impact of social, environmental, psychological and health behaviour factors over time, they can be quite challenging to design, conduct and fund.

In the first paper in this series, Byles, Dobson, Bryson and Brown explore these issues and the challenges that can occur when bringing together a research team to establish a longitudinal study. Flexibility in approach and open communication between members of the multidisciplinary research team and with the funding body emerge as important themes in developing and maintaining a longitudinal study.

Flexibility is also stressed in the second paper, where Warner-Smith, Loxton and Brown discuss the tasks that longitudinal studies involve, the skills necessary for completing those tasks and the organisational structures that support the staff who undertake the tasks. Continuity, succession planning and staff development are all key issues for longitudinal study researchers and staff. Continuity and succession planning are supported by clear record keeping procedures, which allow new researchers, staff members and collaborators easy access to the study methods and procedures. Chojenta, Mooney and Warner-Smith discuss longitudinal study protocols, policies and record keeping in the third paper of this series.

While the first three papers in the series focus on the guiding and supporting structures that facilitate longitudinal investigations, the remaining papers concentrate on the day to day conduct of longitudinal research. Loxton and Young describe the development of initial and follow-up longitudinal surveys in the fourth paper, where the authors discuss striking a balance between desired data, competing research interests and participant burden.

Consideration of participant needs is also an important theme in the subsequent two papers, where Adamson, Young and Byles address recruiting strategies and Adamson and Chojenta examine participant relationships and retention.

Longitudinal studies inevitably collect a large amount of data which increases in volume over time. The management of data that pertain to participant status (cohort data) is the subject of the seventh paper in this series, where Adamson and Graves outline basic principles of database construction that can assist in the management of cohort data. Longitudinal data, the actual research data collected by a study, is the subject of the eighth paper. Here, Graves, Ball and Fraser describe methods for producing clean, accurate and well documented longitudinal datasets. Longitudinal data give rise to a number of analytical issues which are discussed by Young, Powers and Wheway. Among the topics covered by these authors are dealing with attrition, retention, measures of change, and inconsistencies in datasets.

In the penultimate paper of this series, Helman et al. consider the advantages and disadvantages of conducting substudies, also known as nested, side or piggyback studies. This paper also provides procedures that can be applied to the conduct of substudies.

Across all of the papers, a theme of open and documented communication is apparent. In the final paper, communication is directly addressed by Chojenta et al. who report on some of the many ways that longitudinal results might be communicated, including government reports, reports to stakeholders and communications with the general public.

Longitudinal studies have the capacity to answer a wide range of questions about health and wellbeing. Throughout these eleven papers the authors discuss sustainability, communication, documentation, flexibility and adaptability as being key elements in meeting the challenges posed by conducting longitudinal research.
Getting Started: ‘Preparing the ground’ and ‘planting the vines’ for longitudinal research

Julie Ellen Byles
Director, Research Centre for Gender, Health and Ageing, The University of Newcastle, Callaghan, NSW, Australia

Annette Dobson
Professor of Biostatistics, School of Population Health, Faculty of Health Sciences, University of Queensland, St Lucia, QLD, Australia

Lois Bryson
Research Centre for Gender, Health and Ageing, The University of Newcastle, Callaghan, NSW, Australia

Wendy J Brown
School of Human Movement Studies, University of Queensland, St Lucia, QLD, Australia

Abstract
This paper provides a brief overview of some of the practical issues to consider when establishing a longitudinal study. The discussion draws on the experiences of some of the initial investigators of the Australian Longitudinal Study on Women’s Health and sets the scene for the subsequent papers in this series. The investigators reflect on the processes of establishing the research team, conceptualising and planning the research, and gaining and maintaining the funding for the study. Their discussion considers the many disciplinary perspectives that have been integrated into the study, and how these came together. They also highlight some fundamental principles and decisions that must underpin longitudinal studies from the outset.

Keywords: longitudinal research methods; transdisciplinary approach; research planning

The impetus for developing a longitudinal study can come from a variety of sources, but whatever these are, they will almost certainly have implications for the study’s design and development. The Australian Longitudinal Study on Women’s Health (ALSWH), the study that provides the major basis for the papers in this series, was sparked by a government call for a study of women’s health with the goal of informing policy. The project had support among Australian women, with whom the government had consulted widely. While the history of a project and the source of funding clearly do not alter the scientific issues that must be confronted in undertaking longitudinal research, these can, to a greater or lesser extent, affect the focus of a study and the amount of funding can significantly shape its scope. What’s more, issues that are
important at the start of the study may evolve or be overtaken by other issues as the study progresses. Designing a longitudinal project that must maintain alignment with the expectations and needs of a sponsor can also create a set of research imperatives that may differ from those of researchers whose focus often is more exclusively on scientific goals.

In reflecting on our experience of establishing the ALSWH, we are rather drawn to the analogy of creating good wine. Fine wines come from hardy grape vines cultivated in carefully managed vineyards, where the vines have been planted in well prepared and suitable soil, and expertly tended over many years. As with wine production, a longitudinal study requires careful planning and management and, like good wine, longitudinal studies do not yield their best and most valuable product for many years. With this analogy in mind we consider the lessons learnt in establishing our study: in ‘preparing the ground’ for the study by establishing a strong and committed team; in ‘planting the vines’ by determining an appropriate research agenda and measurement plan; and in ‘maintaining the crop’ through securing ongoing funding and the continued commitment of researchers, project staff and participants.

1. PREPARING THE GROUND: ASSEMBLING THE RESEARCH TEAM AND COMBINING PERSPECTIVES

Most longitudinal studies involve a team of researchers with different disciplinary backgrounds. The members of this team and their skills will depend in part on the requirements of the study, on the population of interest, on the theoretical or political paradigm, and on previous relationships, knowledge and trust. The team may include people who know each other and who have worked successfully together in the past, and/or individuals who have come together for the express purpose of the study. Team building is facilitated if at least some of the investigators have worked together in the past.

The team may begin with a core of people who are interested in a certain topic, and grow to include others with particular expertise or interest. For instance, when we began to consider our plans for ALSWH, we started with a core group of people who were interested in women’s health, since that was the topic to be addressed. We quickly expanded this group to include statisticians and epidemiologists, who brought methodological expertise, as well as their interest in public health issues affecting women. The study grew out of a political drive to incorporate women’s concerns into the research, so the core team also included individuals with a particular understanding of women’s policy issues, and psychosocial influences on health, not only through subscribing to a medical or health service model.

Since those early days, the team of investigators and collaborators has expanded to include people with an even wider range of backgrounds and skills, including psychologists, sociologists, nutritionists, other allied health clinicians, medical practitioners, exercise physiologists, economists and so on. The external collaborators involved in the project represent an even broader disciplinary array.

Each member of the team brings their own perspective, priorities, models of theorising and language. The triangulation facilitated by such diverse perspectives can bring strength and depth to the project, but learning to work together across disciplinary divides can be tricky. There are many potential challenges, ranging from topics of interest, to methods of measurement and models for, and types of, analysis.

Finding common ground on the topic of interest may seem to be a relatively easy process. However, at times in the beginning of the ALSWH it appeared that the researchers may have been from different countries, given the diversity of research language spoken in the initial meetings. Words used and understood by one discipline in a specific way were used with contrasting meaning and understanding by others. In the very first meetings called to recruit a team to
develop the ALSWH submission, a division was evident between biomedical models of health and psychosocial models of health. However, the nature of the research was driven by the need to explore women’s health from a broader perspective than that encompassed by either model. Rather than taking an ‘either/or’ approach, the research team worked towards encompassing the best of both models within the framework of the study design.

Bridging these differences requires a great deal of dialogue and translation. It pays to be clear, and to ask the meanings of words. It is necessary to be generous and tolerant of those that are not expert in a particular area so that the epidemiologist can ask what the sociologist means by ‘epistemology’ and the sociologist can ask what the statistician means by ‘degrees of freedom’ without feeling foolish. But beware those terms that you think you understand! Table 1 shows some examples of words that can have different technical and lay meanings.

In addition to linguistic challenges, different research topics may also reflect different value systems among the researchers. For instance, obesity may be seen as a major public health problem from the point of view of the adverse health effects of overweight and its aetiological role in chronic diseases (eg diabetes, cardiovascular disease, asthma etc). However, feminist researchers may see the problem as one of pressuring women to conform to an ideal body shape. A nutritionist may see the problem as one of diet and energy balance, and the exercise physiologist might agree. However, the psychologist may focus on mental health correlates such as depression, the sociologist may focus on life circumstances such as social exclusion, and the health economist may focus on price elasticity and decision economics in their assessment of the problem.

As this discussion on obesity suggests, a range of approaches may be necessary to explore different elements of any research issue, particularly in relation to health. Team members must therefore come to understand other’s disciplinary views and why they matter, not only to the researchers but also to the purposes of the research. It is clear, for example, that to improve women’s health, any focus on obesity without consideration of food and exercise as well as the psychological and social contexts will fail to explore the dimensions of this public health problem, and will overlook possible strategies to address it.

Different disciplines will also have different conventions regarding the extent to which the research must be explicitly hypothesis driven and/or framed by theory. Some researchers may observe and measure patterns in the data, and this is valuable in relation to health behaviours and outcomes which are likely to change over time in response to life changes or government policies. Others will require a theoretical basis to underpin hypothetical models that dictate what data should be collected and how they should be analysed. There may also be debate about which models to use and which elements to emphasise.

For example, while epidemiology, as the study of disease distributions in populations and the factors that influence and determine this distribution, is often the building block of a longitudinal study, perspectives of the factors referred to in its definition vary. While some health psychologists may tend to focus on individual and demographic factors, sociologists place relatively more importance on the social and cultural environments in which people live, work and play. The ALSWH, with its grounding in a social definition of health, has taken what can be termed a ‘social ecological’ approach which combines elements of individual, social and environmental approaches. During the course of the ALSWH, the use of such social ecological models, has become more common.

Longitudinal research is often undertaken by ‘multidisciplinary teams’. In these teams, each researcher has a unique perspective to contribute, and will maintain their disciplinary distinction. In many teams, however, the members learn from each other and eventually develop a broader and more dynamic and comprehensive understanding that crosses disciplinary boundaries. This crossing
of boundaries is referred to as a transdisciplinary approach and can allow for a greater understanding of the complex relationships that can be explored through longitudinal studies. Transdisciplinary approaches can promote new conceptual frameworks which can be used across the individual disciplines involved.

Such transdisciplinary thinking is important when undertaking longitudinal studies of human health. As Albrecht, Freeman and Higginbotham (1998) argue, the total health context includes interdependent systems that span physical, biological, psychological and ecological foundations. For instance, effects that occur within molecular and cellular systems are influenced by (and influence) the person’s behaviour, and their physical and social environment. Consider a biologist’s view of ageing as an accumulation of cellular
damage that supersedes the body’s repair mechanisms and adaptive processes. The clinician may recognize the same process in terms of lack of physiological and cognitive reserve and increased susceptibility to disease, the epidemiologist may relate accelerated ageing to exposure to toxins (such as cigarette smoke, solar radiation), and/or availability of protective agents (such as antioxidants), and the psychologist might consider the importance of personality factors in adaptation to change and response to psychosocial stressors. These different views may all be individually accurate but they are incomplete. If we integrate these different disciplinary perspectives, we can create a more useful and complete conceptual model. For instance, the integration of these different perspectives on ageing may deliver a model in which older people are perceived as vulnerable to physical and psychosocial stressors because of reduced ability to maintain homeostasis, such that seemingly trivial events can lead to serious social, functional and biological problems. Beyond this model, the sociologist’s perspective may represent ageing as a social construction, and the behaviours that accelerate the ageing process as socially determined and including poverty and socioeconomic background. A nutritionist might blame poor health in older age on a poor diet, another researcher might blame poor diet on individual behaviours and choices, while a sociologist might identify the person’s diet and food choices as socially modulated. The truth is that there are too many ‘buts’ for any one discipline to have a monopoly on explaining the full complexity of human health.

In fact, the idea for establishing a study of women’s health was explicitly underpinned by an understanding that health issues require not only a medical focus, but a coalition between a very wide range of physical, social and applied disciplines. This approach has gained almost universal currency in relation to health over recent decades. Hence the project from the outset aimed to meld together researchers from many disciplines and in so doing can be referred to as transdisciplinary. However it remains a contested question as to whether the actual knowledge produced, rather than the methods used, is itself most appropriately termed multidisciplinary, cross-disciplinary, interdisciplinary or transdisciplinary (Aboela et al. 2007; Higginbotham, Albrecht & Connor 2001; Bammer 2001). Like Fairclough (2005), we are cautious about an extreme commitment to postdisciplinarity, recognising that ‘when theoretical concepts or categories and methods from one discipline are incorporated within another, this should be seen as a process [which] requires a careful elaboration of their relationship with existing concepts, categories and methods’ Fairclough (2005: 3).

The answer seems to be that all these terms may be applied to different elements of research, and in some cases this raises complex epistemological issues, which are beyond the scope of the discussion here. Certainly the varied disciplinary backgrounds of the ALSWH researchers have been a great asset in stimulating and sharpening the research in its capacity to deliver insights into women’s health and recognise their potential relevance for policy.

2. PLANTING THE VINES: DEVELOPING THE FIRST SURVEY
Throughout the formation of the research team, the process of defining research goals is inextricably linked to determining the research methods to be used. Arguably, the most commonly used longitudinal research method is the longitudinal survey. Other research methods that focus on, for example, collection of biological samples, will still use surveys to collect demographic and background data.

The first survey of a longitudinal study is critical, not only because it sets the agenda for the duration of the study, but also because it is the first chance to engage the participants. Consequently much time needs to be spent at the start of the study determining the agenda and content of the initial questionnaire. The paper by Loxton and Young (2007), in this issue, provides more
detailed information about survey development for longitudinal studies.

Most longitudinal studies will commence with some broadly defined purpose. For instance, studies of ageing have variously been developed to identify factors associated with ‘ageing well’. The Australian Longitudinal Study of Children (2007) aims to examine ‘the impact of Australia’s unique social and cultural environment on the next generation.’ The vision for the Deployment Health Surveillance Program (2007) is to provide ‘a systematic, prospective and ongoing means of assessing and understanding the health effects of operational deployment on Australian Defense Force personnel’. The ALSWH was designed to identify factors that affect the health of women. In each case the stated aims mark out a very large territory, without a great deal of direction as to how the research team should navigate. The situation is vastly different from other research, such as a randomised controlled trial of a therapeutic drug which might be driven by a single clear and precise hypothesis.

Where to begin? This was the somewhat overwhelming question confronting our original group of investigators. We began by exploring the literature on women’s health issues and by seeking input from a range of experts and stakeholders. Our search was not confined to academic literature but included grey literature and policy documents. Since the initiation of the study largely arose from the Australian National Women’s Health Policy (1989), this was a good place to start.

The Policy outlined a number of areas of importance for women’s health. Major themes identified within the policy included reproductive health, emotional and mental health, violence against women, occupational health and safety, needs of carers, and the health of ageing women. Health system issues were also highlighted. Our review of the literature explored those issues that had been the focus of other longitudinal studies of women’s health such as the United States Nurses’ Health Study (Colditz, Manson & Hankinson 1997), which was designed as a prospective study to gain information on the effects of the oral contraceptive pill and then broadened to risk factors for cancer and cardiovascular disease, the Iowa Women’s Health Study (Steinmetz et al. 1994) which initially focussed on cancer, and the Women’s Health Initiative (2007), which initially focussed on hormonal issues. All these studies had more of a biomedical focus than had been identified by the Australian National Women’s Health Policy, and were designed to collect data on an exposure (eg hormone therapy) in relation to an outcome (eg breast cancer), rather than changes in factors affecting women’s health and associated changes in women’s health over time. The Massachusetts Women’s Health Study (Avis & Mckinlay 1995) was one study that did take a broader psychosocial view of health, but even this did not address the range of issues identified as important within the Australian policy. However, each of these studies did highlight aspects of health and health risks that were worth considering.

Each team member also brought their own priorities and interests, ranging from childbirth to time use, and from medical procedures to psychosocial resilience. Many topics were discussed and debated over many meetings until an organising framework was developed. This framework included a number of themes which have driven the project since that time, albeit with considerable development and evolution, partly in response to changing national priorities, partly due to the interests of the growing team of investigators, and greatly due to the comments and feedback from the study participants. The original themes for the ALSWH are provided below:

• **Use of and satisfaction with health care services**

Includes availability of, access to and utilisation of health services, cultural appropriateness and women’s perceptions of the quality of care provided.

• **Life stages and key events**

Identifies factors that help women to maintain the highest level of health and quality of life.
following key events such as birth of first child, divorce, menopause, widowhood, major illness, moving house, changing job, falls and fractures in the elderly, dementia and bereavement.

- **Time use**
  Provides an indicator of women’s roles in paid and unpaid work and providing care for others. There are questions about social support, overload, independence, leisure, feeling rushed, aspirations and financial resources.

- **Weight and exercise**
  Includes the impact of weight and exercise, dieting and eating disorders on health outcomes. Questions include items about perceptions of body image, as well as general well being and quality of life, dieting and weight change.

- **Violence**
  Explores issues surrounding women’s experiences of physical and sexual abuse and harassment, as well as issues of psychological and social abuse of older women.

These themes were then expanded in our research plan to provide a rationale for their inclusion, some central hypotheses, and the need for specific substudies. Measures relevant to each theme were explored in the methods section of the proposal.

More recently, collaborators involved in the New South Wales 45 and Up Study (2007) have engaged in a similar process. This study involves almost 100 collaborators and is overseen by a Scientific Advisory Committee. In order to develop the research approach, this committee established several theme committees that independently explored possible hypotheses and research questions across domains of health service use, cancer, cardiovascular disease, mental health, and ‘the oldest old’. Each theme committee identified a number of hypotheses that could be addressed over the course of the study. The Scientific Advisory Committee then considered each hypothesis in terms of its novelty and importance, and in terms of its dependence on longitudinal data. Like the ALSWH, this study has also found that the research agenda can be enriched through interaction with key stakeholders and other interest groups.

Likewise the Australian Deployment Health Surveillance Program built on previous studies such as those of veterans in the Vietnam and 1990/91 Gulf wars but emphasises strongly emerging issues from studies conducted in other countries, including post-traumatic stress disorder and multiple non-specific symptoms.

Themes and hypotheses to be addressed provide an essential basis for choosing the measures and methods to be employed in a study, including the sampling frame and inclusion criteria, the mode of data collection (e.g. postal or web-based survey, telephone interview, home visit, clinic visit), and the specific measures to be used.

Having chosen the content areas to be included in the study, the next point of consideration may relate to measurements. Some team members may be more familiar with in-depth open-ended data gathering that allows the participants’ own views to be collected. However, large scale longitudinal studies also require simple, valid and reliable measures that can be easily used to characterise different subgroups and outcomes, and that can be repeated over time. It is also important to use measurement methods that are responsive – that is, they will be able to detect changes that may occur over time.

Some team members may be used to making direct observational measurements, for example of weight and height, whereas others may use measures that require scales with defined psychometric properties and that make indirect assessments around the variable of interest. Everyone will have an appreciation of the importance of the validity of these measures, but will have different interpretation of the dimensions of validity (being seen variously as an abstract truth, conformation to a defined construct, or percentage agreement with some gold standard).

There will also be tensions between those researchers who want complete and rigorous measures and those who are prepared to sacrifice
psychometric rigour for brevity. This tension always exists when choosing measures, but is magnified when many different researchers are competing for space in a questionnaire or interview schedule that is designed to collect a comprehensive range of data across many topics. If, for example, a large amount of time or space is required to measure psychological traits and health beliefs, is there also opportunity for measuring social interaction, health behaviours, symptom profiles, disability ratings and quality of life?

While psychologists are trained in the use of scales formed by combining numerical data based on responses to multi-item questions, and are acquainted with various concepts relating to the psychometric properties of such scales, people from other disciplines may treat the multiple items quite differently. Mathematically trained people should recognise that the sum of several variables will be more reliable (that is, have lower variance) than individual items and be more likely to be approximately normally distributed (from the Central Limit Theorem). However, they may not recognise that this is the rationale for the use of complete scales, rather than using several items from the scales. Clinicians are notoriously keen to reduce the number of questions asked of study participants (to reduce participant burden) without recognising that the validity and reliability of the measurements will thereby be compromised.

Since different measures will produce different types of data (eg discrete or, continuous), researchers from different backgrounds may also have different approaches to analysis. Indeed, at the start of a study, none of the researchers may have worked extensively with longitudinal data. This was true for most people in our team. What is more, this field of research has moved a long way since we started our study – led mostly by advances in computing power and by the fact that researchers have worked on ways to deal with complex multivariate models.

A further issue that is particularly challenging for the design of longitudinal studies is the choice of sample size. It is important to assemble a large enough cohort to represent the population of interest, to allow for attrition over time, and to provide enough diversity to allow differences in exposures and outcomes, and the relationships between them, to be explored. It is helpful to have one or more clear hypotheses at the start of the study, to ensure adequate power to address questions of major importance, even if the study will later expand to address a broader range of issues. Unless the study is initially designed to test specific hypotheses, conventional methods for sample size calculation do not really apply: the initial hypotheses can be diffuse; changeability of measures over time may be difficult to quantify; and the most important questions may actually emerge some time after the study has been established. For the ALSWH we realised that it would be desirable to have a sample at least as large as the US Nurses’ Study but the available budget would not allow this. Therefore we limited the age range of interest (to three 5-year age groups) and then calculated the largest the sample size we could afford! The scientific argument then related to the statistical power that would be available to detect effects of various sizes – some plausible differences would clearly be detectable but others, for example in relation to less common diseases, could not be studied with this design. Further discussion on sampling and sample size considerations is provided by Young, Powers and Wheway (2007) in this series.

3. MAINTAINING THE CROP: FUNDING FOR LONGITUDINAL STUDIES

Large longitudinal studies are expensive and require reliable funding over a long period of time. Identifying and obtaining such funding may be one of the most difficult aspects of the study. In the case of the ALSWH, identifying the initial funding source was relatively straightforward since the study emerged in response to a government tender (however competing for the funding was an intensive and demanding process that required commitment and dedication on the
part of the investigators). Other Australian longitudinal studies have also received substantial government backing from their inception (e.g., Household Income and Labour Dynamics in Australia – HILDA (2007)). The New South Wales 45 and Up Study (2007) obtained its initial funding from a consortium of stakeholders who saw the study as an important resource worthy of investment.

Most Australian studies have had access to sufficient funds at the outset to allow the establishment of the study sample and research infrastructure. However, almost all have had to identify additional, on-going funding sources to allow for the more detailed data collection and analysis that is required to address most research questions. Where this ongoing funding is not secure, the future of a study is obviously highly vulnerable to the success or failure of ongoing funding applications. One researcher tells of how his team made 12 unsuccessful applications to funding bodies for the collection of data to continue their longitudinal study. Thus several years elapsed before follow-up data could be collected.

Other studies have grown from wave to wave on different types of project funding from a variety of sources that might reflect different priorities and approaches. Mostly these sources will not fund ‘a longitudinal study’ in general, but will fund the testing of different hypotheses using a longitudinal design. The ALSWH has accessed additional funding for many specific projects that build on the basics of the longitudinal design, including investigations into eating disorders, road traffic accidents, weight management, alcohol, asthma, violence against women, diabetes care, caring for older people, processes of retirement, incontinence, sleeping difficulty and sleeping medication use (to name a few).

Such ‘add on’ studies and in-depth analyses are essential if the full value of the information collected in longitudinal studies is to be realised. Many longitudinal studies, even those designed to answer a specific initial hypothesis, have the potential to yield information on a much greater range of topics. For instance, the Blue Mountains Eye Study (Mitchell et al. 1995) has also provided information on nutrition, driving, and smoking, among other topics. The untapped potential of many existing Australian longitudinal studies of ageing was recently recognised by the NHMRC Ageing Well Ageing Productively program that provided substantial funding to enable integration of data and findings and further analyses across nine longitudinal studies of ageing.

In addition to securing funding from traditional ‘research’ sources, the ALSWH researchers have from time to time considered the issue of corporate sponsorship. This is one issue which engendered markedly different responses from researchers with different disciplinary perspectives and values. While everyone, and their academic institutions, would be against accepting money from the tobacco industry (where the product from the profits is bad for everyone’s health) what about accepting funding from makers of alcohol? In moderation alcohol lowers the risk of heart disease but in excess it is harmful. What about the pharmaceutical industry which saves millions of lives each year but is seen by some as dangerous and exploitive in certain situations? Soft drink manufacturers would be an unacceptable source of funding to nutritionists and meat producers would be unacceptable to vegetarians. In the ALSWH there was even debate about the merits of seeking sponsorship from manufacturers of detergents and cleaning products, and from makers of cosmetics, because of the perceived stereotyping from portraying women in association with these products! Such differences of opinion need to be treated respectfully, and in the context of how urgently the study needs funds in order to continue.

4. REAPING THE HARVEST: COMMUNICATING THE RESULTS

For the ALSWH, the research plan particularly needed to consider the potential end users of the study findings, including policy makers. We needed to consider what audiences might make use of
the findings and how the results might be communicated and made available to them. Our original planning for the study took account not only of how we would report to our funding bodies, and how we would communicate with fellow scientists through peer reviewed publications, but also dissemination of the findings more widely through mass media. In fact, communication of the findings in these ways, and through seminars, has proved to be a major ongoing activity of the study. Further discussion of our communications strategy is provided in the final paper of this series (Chojenta, Byles, Loxton & Mooney 2007).

5. Further considerations for starting longitudinal research

Time is the most important feature of a longitudinal study. If an ‘exposure’ is the cause of an ‘effect’ then it must be observed to have occurred before the effect does – and this crucial time order for causation can only be unambiguously determined using a well-conducted, prospective, longitudinal study. Thus the first issue in establishing a longitudinal study is to identify mechanisms to collect data on the same people over time. This means that participants must be selected and invited to be followed for the duration of the study. Ways of keeping in contact with them need to be established, and thought given to how to maintain their interest and commitment to participation in the study (see Adamson & Chojenta 2007). The study must establish a very close, respectful and enduring relationship with these people.

Researchers will sometimes take some time away from the study to pursue other interests, but the cohort of participants is irreplaceable.

Measures that will be of importance in the future must also be identified and collected at the outset. Most researchers on longitudinal studies would no doubt desire a time machine – either to fast forward to when their outcome data will be available and the study can achieve its outcomes, or to rewind to add to or correct errors in their earlier measurements. But in the absence of such an invention, all that can be done is to under-

stand as well as possible the issue of time.

Time allows researchers to gather data on incidence of health outcomes (ie new cases) rather than prevalence (all cases in the population), and time allows researchers to establish temporal relationships between suspected exposures and outcomes. Time is also the ingredient that is most lacking at the outset of the study, so that it is sometimes argued that the most important longitudinal studies are birth cohorts which actually recruit mothers before the study subjects are born (although even birth cohorts may lack information on parental exposures). More generally it is desirable to collect data before key exposures occur (such as from defence forces personnel about to be deployed to a war zone).

When research questions and hypotheses do not relate to time (such as cross-sectional associations between variables), a longitudinal study is not the right vehicle for studying these phenomena. In the ALSWH we are often approached to conduct substudies that really only need a sample of women and do not have a time component – these requests are rejected because it would be unethical to treat the ALSWH cohorts as a ‘convenience’ sample and to increase participant burden in this way.

Considering the ethical aspects of the study, approval from the relevant Human Research Ethics Committee/s is a requirement of all research involving humans, and it is likely that approval might need to be sought from more than one institution when research team members are based at more than one university. This can take considerable time and needs to be allowed for in the planning process. The National Health and Medical Research Council provide guidelines for the ethical conduct of research (2007) and other documents that offer information and advice about ethical procedures. The ALSWH has found the assistance of university ethics officers to be highly valuable in developing ethical procedures, particularly when including survey questions about sensitive issues (eg violence, illicit drug use).
Arguably the greatest difference in ethical implications between longitudinal and cross-sectional studies is the need to retain contact details of participants for long periods of time, perhaps even for the participant’s lifetime. The need for secure storage and access procedures for these details is a prime consideration for longitudinal studies (discussed further by Adamson & Graves (2007) in this issue).

Over time, codes of ethical conduct will change, and approvals may need to be varied and updated. The procedures guiding ethical research with humans were, in fact, revised by the NHMRC in 2007. With each revision of ethical practice guidelines, it is incumbent on the research team to ensure that the longitudinal study procedures and materials remain ethically sound. Likewise, every change in longitudinal study procedure requires ethical clearance before being implemented.

CONCLUDING REMARKS

Longitudinal studies are like a good wine which requires hard work in selecting and planting the vines, care in harvesting the grapes at the right time, meticulous attention to processing the yield, and protection of the juice while it comes of age. Much is to be gained from working towards a transdisciplinary team approach, by blending concepts from differing disciplinary perspectives, while not losing the methodological strengths provided by the individual scientific disciplines. However, it is important that all the researchers share the same broad aim. In the case of the ALSWH, our aim is to impact on public health policy and practice for the ultimate benefit of all women.

As this series of papers draws largely on our experiences in establishing and maintaining the ALSWH, it cannot provide all the necessary training in longitudinal research design and management. Many good books and papers are available for that purpose. Rather, what we aim to provide is some detailed insight into some of the practical issues related to running large longitudinal studies, and how to get the most return from the substantial investments made by funding bodies, collaborators and research participants. Much of the processes and insights described were developed by the project staff. The project staff have been instrumental in developing efficient procedures, building positive relationships with participants, and maintaining the security and integrity of the longitudinal data. The paper by Warner-Smith, Loxton & Brown 2007 in this issue defines the professional skills necessary to complete these tasks, and describes principles for hiring of staff and investing in staff development.

Acknowledgement

The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women’s Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

References


CALLS FOR PAPERS

A special issue of *Journal of Family Studies* – volume 15/2 (August 2009)

**Family–Work Balance**, Edited by David de Vaus

Deadline for Papers: 31st July 2008

1. **Demographics of work and family**
   a. Trends in workforce participation of parents with young children. Change over time and across the life course. Who works and who doesn’t?
   b. Fertility and workplace policy re families

2. **Use of formal and informal child care**

3. **Influence of paid work on family life**

4. **Economics of workforce participation decisions: incl disincentives**

5. **The changing organisation of work and impact on family**
   a. Extent of workplace flexibility for family issues
   b. Unsocial hours
   c. Impact of industrial changes on family life

6. **Attitudes and preferences**
   a. Attitudes about work and family issues
   b. Reasons for mothers working
   c. Work preferences of mothers with young children

7. **Domestic division of labour with arrival of children**

8. **Work and family in later life**
   a. Elder care
   b. Synchronised retirement

9. **Role of law, government and other regulations in the work family balance**
   a. Industrial legislation, determinations, child care and other benefits

Authors are invited to discuss their proposal with the editor in advance of manuscript submission:

**Professor David de Vaus**
Faculty of Humanities & Social Sciences, La Trobe University, Victoria 3086, Australia
Tel. +61 (03) 94792844

Please send papers or expressions of interest in the first instance to the Assistant Editor: l.morrison@latrobe.edu.au

Author guidelines can be found at: http://jfs.e-contentmanagement.com/page/16
Human resources for longitudinal studies: Matching people to skills and tasks

PENNY WARNER-SMITH
Research Centre for Gender, Health and Ageing, The University of Newcastle, Callaghan NSW, Australia

DEBORAH LOXTON
Research Centre for Gender, Health and Ageing, The University of Newcastle, Callaghan NSW, Australia

WENDY J BROWN
School of Human Movement Studies, University of Queensland, St Lucia, QLD, Australia

ABSTRACT
This paper describes the practical tasks that longitudinal studies involve, the skills necessary to complete those tasks, and the organisational issues that are pertinent to conducting longitudinal research. We first focus on the decision about whether to conduct the study in-house, or to outsource part or all of the work. We discuss the desirable qualities of the people responsible for carrying out the project tasks, and the importance of creatively matching the work to the skills and experience of the people in the organisation. The critical issues of continuity and succession planning are then addressed before we finish with a brief discussion of organisational structure. Throughout the paper, examples from the Australian Longitudinal Study on Women’s Health (ALSWH) are used to illustrate the issues at hand. The paper is one of a suite of eleven papers which address the practicalities of running a longitudinal study.

Keywords: longitudinal studies; human resources

One of the most common questions we are asked about running a longitudinal study is, ‘Who does what?’ Conducting a longitudinal study involves many tasks undertaken by people in a variety of positions, some of which are more obvious than others. For example, it is easy to see that someone will be needed to manage participant databases and research data, but perhaps less obvious that both ‘higher level’ (e.g. financial and human resource management) and ‘lower level’ (e.g. filing) tasks are vitally important.

In describing the tasks that longitudinal studies involve, the skills necessary to complete those tasks, and the organisational issues that are pertinent to conducting longitudinal research, this paper contributes to an area in which there is a dearth of literature. While the paper can be read on its own, it is one component of a suite of eleven papers which address the practicalities of running a longitudinal study. It is most fruitfully read in conjunction with the companion papers which cover a variety of topics ranging from set-
setting up a long study, to statistical issues, to questions of protocols and record-keeping.

Specifically, in this paper we first consider the question of who will do the work, focusing on the decision about whether to conduct the study in-house, or to contract part or all of the work to one of the increasing number of specialist data processing companies. We then consider the particular tasks that have to be done in running a long-term research project, and the desirable qualities of the people responsible for carrying them out. We emphasise the importance of creatively matching the work to the skills and experience of the people in the organisation. The critical issue of continuity and succession planning is then addressed before we finish the paper with a brief discussion of organisational structure. Throughout the paper we use examples from the Australian Longitudinal Study on Women’s Health (ALSWH) to illustrate the issues at hand.

Two main groups of people are involved in large research projects such as longitudinal studies. The composition and gathering together of the academic research team has already been discussed (see Byles et al. (2007) in this issue). The other important group is the staff who manage and conduct the clerical, administrative and data management tasks. While many research projects involve these tasks, longitudinal studies require longer term staffing solutions than short-term or ‘one-off’ project based research. The types of task to be performed are also unusual, so that it might be difficult to find people with specific experience in longitudinal projects.

1. WHO WILL DO THE WORK?

In considering human resource issues it is important to consider whether it is necessary to have ‘in-house’ staff who have the skills for all aspects of the project, or whether it might be more economical to ‘outsource’ some tasks to specialists. Over the past ten years there has been a trend towards outsourcing or subcontracting aspects of research projects. This has been particularly true for the printing and distribution of materials (eg surveys), data collection (eg interviews), data processing (eg inputting survey data) and data management (eg production of datasets, maintenance of participant lists). It is therefore likely that the first decision that any new longitudinal study group will face with regard to staffing the project is whether any of the tasks will be outsourced, and if so, which ones. This is a tall order for a new project, where the tasks themselves might not be fully developed. Some guidance can be found in the following section.

### Table 1: In-house and Outsourced Processing: A Comparison

<table>
<thead>
<tr>
<th></th>
<th>In-house</th>
<th>Outsourced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control of processes and data</td>
<td>Full control is maintained</td>
<td>Control is limited by company policies</td>
</tr>
<tr>
<td>Quality control</td>
<td>Quality control is assured</td>
<td>Quality control standards are determined by others</td>
</tr>
<tr>
<td>Budgeting</td>
<td>Budget is more flexible</td>
<td>Less flexibility in budget</td>
</tr>
<tr>
<td>Participant contact</td>
<td>Participant contact is standardised</td>
<td>Participant contact is dictated by company</td>
</tr>
<tr>
<td>Ethical issues</td>
<td>Regulated by research ethics committees and standards</td>
<td>Ethical precautions harder to oversee</td>
</tr>
<tr>
<td>Student employment</td>
<td>Provides employment opportunities</td>
<td>No student employment opportunities</td>
</tr>
<tr>
<td>Intrinsic interest of the work</td>
<td>Some work is repetitive and not challenging</td>
<td>Less interesting tasks can be avoided</td>
</tr>
<tr>
<td>Equipment</td>
<td>Equipment purchases may be costly</td>
<td>Equipment purchases unnecessary</td>
</tr>
<tr>
<td>Cost efficiency</td>
<td>Some tasks may be less cost-effective (eg bulk mailouts)</td>
<td>More cost-efficiency for some tasks (eg bulk mailouts)</td>
</tr>
</tbody>
</table>
Outsourcing has both advantages and disadvantages (see Table 1). For example, contracted companies can provide good ideas about data entry and processing. However, the obvious disadvantage is that the suggested methods will be those that the company uses and may not necessarily be tailored to suit individual projects. It is therefore crucial to have a detailed plan in place before shopping around for contractors (see Table 2). It is also important to undertake site visits. Different companies will use different processes for mailouts, receipt of surveys, data entry and delivery. Site visits allow an accurate comparison of these often complex processes to be made. In addition, site visits allow for the foundation of a positive relationship to be laid between the project and the company. For example, being able to put a face to a name when discussing the job later via email or by phone saves time and is likely to lead to a better product.

The ALSWH has adopted a flexible approach to outsourcing. Where a task can be efficiently and economically managed by an external company, and where there is no obvious disadvantage, or loss of advantage, to sending the work off-site, then work will generally be outsourced. In ALSWH we initially explored the possibilities of undertaking most of the tasks in-house, but decided that greater efficiencies could be achieved by outsourcing the printing, mail out and data entry. Initially, surveys were logged ‘in-house’ but with the advent of newer technologies, we have moved to outsourcing receipt of returned surveys as well as data entry. Although outsourcing can save time for in-house staff, allowances must be made for the time that will be required to liaise with contracted companies, set up contracts and design the processes that facilitate outsourcing.

One task that has not been outsourced in the ALWSH is the maintenance of the participant (name and address) database. We decided to maintain this database ‘in-house’ for the following reasons:

- **Confidentiality:** Ensuring that name and address details of participants would remain confidential was particularly difficult. At two of the companies visited in the course of deciding on a subcontractor, ALSWH staff were shown confidential documents that were currently being processed. These breaches of confidentiality by the companies concerned led to their immediate ineligibility to process ALSWH surveys and data. These experiences further underscore the point that site visits are essential before signing any agreements with subcontractors.
- **Cost:** The cost of moving the participant databases off-site was prohibitive when compared to the cost of keeping this task in-house. In addition, where in-house costs can be adjusted if necessary to account for budgetary constraints, outsourced costs are set by the contracted company.
- **Flexibility:** Holding the participant database in-house means that ALSWH staff can add information or fields to the database at will. For example, dates of death can be added, or variables manipulated to give new information.

<table>
<thead>
<tr>
<th>Table 2: Tips for subcontracting tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preparation</strong></td>
</tr>
<tr>
<td>Have a very clear plan for the work to be outsourced</td>
</tr>
<tr>
<td>Do your homework and allow time to determine the relative merits of in-house versus outsourced processing</td>
</tr>
<tr>
<td>Know how much it would cost to perform the service in-house</td>
</tr>
<tr>
<td>Become familiar with jargon (eg printer’s terms)</td>
</tr>
<tr>
<td>With assistance from your institution’s legal or business office, create a detailed tender document</td>
</tr>
</tbody>
</table>
(eg changing the ways in which data are coded for geographical area). Outsourcing this database would mean adhering to the protocols of the company concerned, which would make it more difficult to make modifications to the database as required.

• **Personal contact:** There were concerns among ALSWH staff that turning over participant contact to a company would depersonalise the project, especially if maintenance of the participant database involved tracking participants (ie finding participants who have moved and not provided the ALSWH with new contact details). For example, ALSWH staff currently check personal details with participants by telephone and, while doing this, are able to answer any enquiries about the project that participants might ask. An external company would not have the comprehensive knowledge to be able to answer these questions. Arguably, this personal element has helped to maintain participant goodwill and trust and has contributed to the study’s relatively high retention rate.

• **Efficiency:** Over the past ten years ALSWH staff have perfected a range of tracking techniques (discussed in Adamson & Chojenta in this issue). No company was found that had a system that equalled the efficiency and success rate of the ALSWH team.

• **Opportunities for students:** For some time ALSWH has employed a number of university students as Project Assistants (PAs; discussed later in this paper). While working as an ALSWH PA provides the students with extra income and thus helps them to complete their education, it also enhances their career prospects by providing practical research experience. Although the ALSWH does not exist as a training ground for students, offering these opportunities is in keeping with philosophies of higher learning.

2. **What has to be done?**

Regardless of whether or not work is outsourced, the tasks that need to be completed remain similar. For example, data have to be collected, entered, processed, and then produced in a usable format. Longitudinal studies also involve specialised jobs, such as tracking participants, developing and maintaining participant contact databases, maintaining contact with participants, and ensuring longitudinal data integrity.

This section describes the tasks that have to be performed in order to conduct a longitudinal study. It is important to maintain some flexibility in terms of matching people and their skills with the tasks to be completed, so the tasks that follow are not position descriptions. Many tasks are completed by groups of people working together – team work and problem solving are therefore key attributes to look for when selecting staff. Moreover, as longitudinal studies have peak periods and quieter times, staff are likely to undertake different tasks at different times in the data collection cycle, so that ‘permanent’ staff may need more assistance from casual or part-time staff at specific stages of the project.

### Project management or co-ordination

The major tasks involved in managing a longitudinal study are human and financial resource management, overseeing research processes, and managing internal and external communication.

**Management of human resources** includes:
- Recruiting and retaining staff
- Developing position descriptions
- Matching people to skills and tasks
- Training staff
- Conducting appraisals and performance reviews
- Ensuring occupational health and safety
- Implementing institutional human resource policies.

**Management of financial resources** includes:
- Developing and monitoring budgets
- Managing accounts
- Planning for longer term finances.

A considerable amount of management time is also spent in overseeing research processes,
such as data management (eg assessing new data collection techniques, the viability of outsourcing and selection of subcontractors), participant contact and related duty of care responsibilities, seeking and maintaining ethics approvals, and managing sub-studies. Management time is also spent in aspects of communication, both internally and externally, and in overseeing the production of ‘deliverables’ (eg reports) that are outside the ‘academic’ sphere of publishing in academic journals. Good communication skills are therefore vital, both for negotiating contract details as well as responding to enquiries from a range of stakeholders.

**Data management**

As will become apparent, data management of a longitudinal study involves a diverse range of skills (see Adamson & Graves; and Graves, Ball & Fraser in this issue). Specifically, in addition to working with the researchers and project manager to oversee data collection, there are two major data management tasks:

- **Managing the participant database**
  This involves maintaining the participant database and producing reports of participant status (as required by researchers, statisticians, and other staff), producing mailing lists (eg for newsletters and sub-studies), and designing and implementing security measures to protect confidentiality.

- **Managing the ‘research’ data base**
  This includes tasks such as assisting with survey design, developing systems for data entry, screening, checking and cleaning raw data following data entry, monitoring response rates, creating scale and aggregate scores, producing frequency tables and other statistical resources, as well as developing, formatting and maintaining all the datasets.

**Statistics**

Longitudinal studies always involve statistics. While data analyses can often be conducted by outside parties, the conduct of some statistical analyses is an integral part of running a longitudinal study. Statistical tasks include:

- Assessing and imputing missing data
- Examining the nature and impact of sample attrition
- Developing and recommending analytical techniques
- Conducting psychometric analyses of survey data
- Preparing reports for funding bodies
- Liaising with study collaborators on statistical techniques
- Working with other researchers on data analysis for publications
- Developing processes for linkage with other datasets.

**Communication**

Effective communication is a very important factor in a longitudinal study (see Chojenta, Byles, Loxton & Mooney in this issue). In this context, ‘communication’ covers a wide variety of tasks, ranging from those that are integral to the management of people, finances and data (which we largely consider as ‘internal’ communication), to public relations and communication with participants and stakeholders (which we largely consider as ‘external’ communication). During the first ten years of the ALSWH, technological advances, including the internet, mobile phones, tele- and video-conferencing, have facilitated both internal and external communication.

Some of the specific communication tasks which are important in longitudinal studies are:

- **Internal communication**
  - Keeping minutes of meetings and ensuring that the reasons for important research decisions are documented. This is extremely important – we have lost count of the number of times we have been asked ‘why did you do that?’
  - Ensuring that there is good communication between academic researchers and the
people ‘at the coal face’ who are collecting and managing the data
– Ensuring good communication in work groups.

• External communication
  – Web page management including designing and maintenance of the web page
  – Maintenance of databases of collaborators and researchers, media contacts, stakeholders (eg people who might receive annual reports)
  – Design, writing and distribution of informal project updates, formal reports, participant newsletters etc.

Tracking
Participant tracking is a very specific and important task in longitudinal studies, and is covered in more detail in Adamson & Chojenta in this issue. The work involves:
• Accessing databases to obtain information
• Telephoning participants or their friends/relatives
• Tracing participant movements using a variety of information sources
• Following paper trails
• Updating databases.

Administrative duties
While everyone working on the project undertakes a number of administrative duties, there is often a need for someone to provide assistance with everyday administrative tasks. In some universities more specialised administrative tasks are undertaken centrally (eg legal services, accounting, and human resource services). There will however always be a need for someone in the research group to provide day-to-day assistance with a wide range of administrative tasks such as:
• Account keeping, including processing accounts payable and receivable, monitoring expenditure and income, producing summary reports etc
• Booking teleconferences, travel and accommodation for research meetings
• Processing incoming mail and responding to enquiries
• Entering data and updating databases
• Processing and logging incoming mail (perhaps including surveys)
• Taking minutes of meetings, recording study procedures
• Assisting with the organisation of printing and mailouts (eg newsletters, reports)
• Filing, filing, filing...

Specialist and technical duties
All research studies involve a number of tasks which require specialist skills. For example, collection of biological samples could require qualified staff to collect samples, and laboratory assistants to analyse and store samples. Survey studies require the assistance of people trained in psychometrics. More recently, studies have involved the use of handheld computers (or PDAs) by participants (Soupourmas et al. 2005), pedometers and other items of technical equipment. It is beyond the scope of this paper to detail the many specialist duties that might be required for a given study. However, researchers can assess their potential needs by discussing their methodologies with other researchers who have conducted similar research, specialists in the field (eg nurses, IT support staff), or their institution’s research office.

3. Matching tasks to people
The number of people required to undertake the work described in the previous section is finally determined by the size of the project and available funding. In making decisions about the composition of the original ALSWH ‘team’ we considered not only the tasks, and the skills required to complete them, but the ability of applicants to ‘problem solve’ and work as part of an integrated team. This was important because many of the tasks were being undertaken for the first time, and there were not tried and tested
protocols to work from. Over the years we have often found that enthusiastic and ‘willing to learn’ applicants are more valuable in the long term than people with more ‘conventional’ skills in a particular area, as the needs of the project are continually developing and changing over time.

As there are no hard and fast rules about people and positions, (or who should do what), flexibility is a key staff selection criterion. This section describes some of the other essential and desirable qualities of personnel involved in different aspects of a longitudinal study.

**Project management or co-ordination**

As indicated above, the person or people responsible for managing or co-ordinating the project should be good communicators because they are the conduit through which the principal researchers’ operational decisions are implemented. Although managers and co-ordinators are usually responsible for managing staff, as well as for maintaining high levels of internal and external communication, they are also responsible to the researchers and funding bodies. They therefore need high level management, organisational and communication skills, and the ability to implement the decisions of others and to have input into decision-making on behalf of others.

Experience in both human and financial resource management is desirable. However, as systems vary widely it is likely that some training will be offered to managers to breach any gaps in knowledge of specific systems. The task of overseeing a longitudinal study’s operation is complex and having good organisational skills is essential.

Importantly, the people in ALSWH management positions have always had the dual function of being managers and research academics. In addition to their managerial roles, they have the usual responsibilities of academics, such as developing their own research programs, publishing in peer-reviewed journals and supervising postgraduate students. The advantages of this are that the managers have a thorough understanding of the research process and methods, familiarity with ethics and ethical procedures, and the ability to apply for supplementary funding through grant schemes.

**Data management**

Data managers require experience and knowledge of all of the database programs that the project will use. Because information technology is continually evolving, data managers should also be willing to undertake further training when required. For a senior position in data management, at least two years experience in a similar position is desirable. The participant and collected data databases are the two most important assets of a longitudinal study. The people who design and maintain these assets need to be reliable, trustworthy and have the ability to work both independently and as an integral part of a team.

**Statistical expertise**

Statisticians obviously need experience in manipulating large and complex data sets as they are not only responsible for conducting statistical analyses, but also for ascertaining the quality and reliability of the data sets. It is desirable that statisticians have experience in longitudinal analyses but as this is an emerging field subject to constant change, and analysis software is under continual refinement and development, it is just as important that they be ready to undertake further training. Statisticians might also be called upon to develop new techniques for analysis and data linkage, so they should be willing to explore solutions to complex challenges. As with most positions on a large scale study, statisticians should be able to work independently and as members of a team.

**Communications/public relations**

It is perhaps unnecessary to point out that the person who undertakes the main communication tasks for the study should have excellent written and oral communication skills, and preferably have a qualification in media studies or a related
field. Experience in web design, database management, graphic design and the production of documents is desirable for a communications officer. Some of these tasks could be conducted at an institutional level, in which case the project manager and/or administrative workers could oversee those tasks.

**Tracking**
Tracking is an unusual job, so it is unlikely that people who have experience in tracking will be available for positions on a new study. Although the tasks are specialised they are easily learned and undertaken. Desirable qualities for trackers include a willingness to learn, a friendly telephone manner, hobbies or interests that include puzzle solving, and the ability to work as part of a team. After-hours availability is also important, as many participants are unlikely to be available during business hours. At the ALSWH, tracking is conducted mainly by university students on casual or part-time contracts, and this has the added advantage of giving the students research work experience. The trackers are trained and supervised by more senior personnel.

**Administrative duties**
The administrative duties required for a longitudinal study do not differ significantly from those of a more short term operation. However, when a project is long term, consideration needs to be given to preparing for staff turnover, which will be discussed later in this paper. The ALSWH has always had an administrative officer who has provided assistance with a wide range of administrative tasks.

**Research**
Most of the skills identified in this section relate to the actual collection of data. There is however also the very significant issue of who will do the data analyses and write up the findings of the research. In the case of the ALSWH, government funding is only sufficient to collect the data and prepare specific reports. As is the case for most other longitudinal studies, the ALSWH is largely dependent on academic researchers who are supported by universities to analyse data and write up the research findings. There are now investigators and collaborators at a wide range of institutions in Australia and overseas.

This situation brings with it a series of problematic issues concerning who has access to the data. Protocols have been developed for providing the data, ensuring confidentiality, attempting to maintain some control over the content of academic journal articles, and trying to make sure that people don't unknowingly duplicate the work of others. These issues are discussed in more detail in Chojenta, Mooney & Warner-Smith in this issue.

**Students**
Longitudinal studies also provide a wonderful base for student research, being a source of existing data as well as a potential base for further research. Most longitudinal studies can provide access to supervisors who are familiar with the project, and the general staff can offer practical support to students.

In the case of the ALSWH, only a small number of academic researchers are physically based in the two principal study locations. Consequently, most students (more than three-quarters of those who conducted their research between 1996 and 2006) are based in other academic schools or departments, as well as in other universities. Co-supervision of these students is usually organised with academics associated with the study.

Between 1996 and 2006, 24 research students completed their theses using data from the ALSWH and at the time of writing a further 15 postgraduate students are enrolled. This brings considerable benefits to the study, in terms of increased use of data, dissemination and output of results, and awareness of the project in the wider community. This is particularly important for longitudinal projects where the volume of data is continually increasing and there is a need...
to publicise results when the issues are most pertinent, rather than waiting for existing researchers to finish current projects. Given that the principal funding for the study must be allocated to data collection, the ALSWH has had limited resources to support students and therefore student researchers are usually funded through external scholarships and project grants.

Staff development
We have already pointed to the need for further training, particularly for people such as data managers and statisticians. As the ALSWH project has evolved there has been a continuing need to provide opportunities for staff to learn new skills, not only for their own personal development, but because emerging technologies become available which can improve efficiencies of data collection and management. For example, voice recognition software has enabled entry of qualitative data into a searchable database. Previously, these data were hand entered by data entry operators.

As significant resources, particularly in regard to time, are needed for staff training, budgets should take account of the costs of both the training and the cost to work outputs of staff absences during training. For example, a recent training course was attended by three ALSWH staff for three days. Although the course costs were minimal, the total time lost in terms of productivity was 63 hours.

4. Continuity and succession planning
At one of the first presentations about the ALSWH in 1998, Professor Bryson (one of the founding investigators) was taken aback when asked what was being done about succession planning. While it may seem obvious in hindsight, at that time this was something that none of the researchers had considered. The ALSWH is planned to run for at least 20 years: in that time investigators and staff might move on to other projects or retire. Who will replace the people as they move on, and what processes should be put in place to make these inevitable transitions problem free?

One strategy which has occurred naturally is that several previous research students have already become project ‘investigators’, ensuring strong corporate memory and continuity between one generation of researchers and the next.

Continuity of staff
As time goes on, some of the original investigators and staff will inevitably move on or retire, and new people will become connected with the study. ‘New blood’ potentially brings changes in research focus, and new people suggest new research questions which reflect their area of interest and expertise. The influx of new staff and investigators over time allows for new synergies to be developed, keeps energy levels high and can create a dynamic atmosphere for continuing high quality research.

However, it is impossible to stress too strongly the importance in a longitudinal study of continuity of staff. One barrier to continuity of staff is uncertainty of funding. People worry about their financial security, and instability in funding can take its toll on staff morale, so staff move on. When good people are recruited, the longer they remain with the project the more valuable they become. They have a stock of institutional memory which is invaluable. For example, one of the long term ALSWH statisticians can for the most part recall which questions have been repeated in each survey and which have not, so thorough is her knowledge of the questionnaires. The senior project assistant knows many participants individually because she has been answering their phone calls and responding to their inquiries since the ALSWH began. Over time, individuals develop contacts and tap into relevant networks, or they may establish new ones, which benefit the project. Perhaps most importantly, maintaining institutional memory allows for a full assessment of any proposals to change procedures to be undertaken, particularly if these ‘new ideas’ have been tried and found wanting during the past.
**Multiple keepers of the secrets**

In a longitudinal study it is extremely important that more than one person holds 'the keys' to any aspect of the project. As the complexities of the project increase, it could be potentially disastrous for a key person to 'fall under a bus'. Indeed, when the original group of ALSWH researchers planned to attend a public health conference in Perth to present some of the initial findings from the study, it was suggested that they should not all travel on the same plane! During the first review of the project it was even suggested that linkage identification numbers be kept in a Swiss bank vault!

**Job manuals**

In order to maintain continuity, a tool that has been implemented at the ALSWH is the 'Job Manual'. These manuals provide a detailed description of the major components of a position or specified task. Short-term studies might benefit from having job manuals; however although study protocols are necessary to ensure good standards, job manuals as such are not essential to the successful completion of the research project. By contrast, longitudinal studies might run literally for a lifetime, and it is highly unlikely that the people who are hired at the outset will still be there, in the same positions, when the research is concluded. Job or task manuals can help to smooth the transition from one staff member to the next.

At the ALSWH, each staff member keeps a job manual. It is a requirement, assessed annually, that each staff member keeps their job manual up to date. Manuals are generally divided into tasks, so if a task is moved from one position to another, for example if an administrative assistant takes on an aspect of database management, the relevant section is moved from one job manual to the other.

It is also useful to keep sections of job manuals that have become redundant. When a task changes significantly, for example if new software is introduced, old tasks become outmoded and will be removed from the manual. However, at some future point, it might be useful to know how a particular process used to work. By archiving redundant sections, institutional memory for processes is preserved.

5. **Organisational structure**

A formal organisational structure will assist with communication between investigators, staff, the funding body and other stakeholders and will also facilitate succession planning. Organisational structures need to take into account the roles and responsibilities of staff members, the supervisory roles of senior staff, the leadership roles of investigators, the relationship between the longitudinal study and its funding body and other stakeholders, and provide opportunities for professional development and promotion.

A key factor to be considered when designing the structure is the aims of the study. These include meeting contracted deliverables and investigating specific research areas or themes. These aims are met through a set of objectives such as producing data sets, conducting analyses and substudies, and designing media for research dissemination (eg web pages, reports, publications). The structure then must support the processes that need to occur for the aims and objectives to be met.

The difference between a longitudinal study and a shorter term research study is that some of the issues focused on initially are likely to change not only with the inevitable changes in the social environment but also as the research evolves, as the interests of investigators change, as new investigators come on board, and as the needs of the funding body alter. The organisational structure should be designed to support both goal achievement and changing goals.

The simplified depiction of the current ALSWH organisational structure is presented in Figure 1. The funding body in consultation with the study investigators determine the contract deliverables (currently once every three years). Research themes are determined by the Steering Committee in consultation with investigators, the Project Advisory Committee and the funding body. Themes are reviewed by the Steering Committee on a yearly basis. The Steering Committee comprises investiga-
tors and senior project staff who oversee the processes involved in meeting goals. Assisting in the implementation of these processes are the Publications, Substudies and Analyses Committee (PSA) and the Data Management Group (DMG). The PSA, which is a subcommittee of the Steering Committee and includes Steering Committee members and advisors (on-staff statisticians) oversees the analyses and substudies that are undertaken, and also monitors output. The DMG comprises representatives from the Steering Committee, project and data managers and statisticians. This group is responsible for developing methods for managing the data (both participant and project data) and making recommendations for statistical analyses. The DMG also monitors attrition, missing data, and anomalies in the data sets.

No single organisational structure will fit all longitudinal research studies well. As with matching tasks and people, flexibility in organisational structure is essential if the project is to adapt to the changing research and social environments over the course of a longitudinal study.

The development of the ALSWH organisational structure
When the ALSWH started in 1995 the organisational structure included a project manager, data manager, statistician, three research assistants and a secretary at the University of Newcastle. This group worked with the project investigators to develop the first wave of surveys, which were conducted in 1996.

Once the ‘main’ surveys had been despatched (simultaneously to women in three age groups who would become the three main cohorts), there was a need to employ many casual staff to deal with the vast quantities of data and never ending telephone calls from prospective participants. As subsequent surveys and substudies were developed, the complexity of tasks grew, and with this the need for staff training and the natural progression of staff into more senior supervisory and management positions.

While the ALSWH gains strength from being conducted across institutions and by engaging researchers from many organisations, this does require that the organisational structure is built to accommodate this added complexity. This structure underscores the need for very good communication, which is achieved through monthly steering committee and data management teleconferences, as well as face-to-face meetings.

In Newcastle the Study is now situated in the Research Centre for Gender, Health and Ageing (University of Newcastle), and the Centre Director is now the Co-director of the ALSWH. In Brisbane, the Study is based in the School of Population Health (University of Queensland), where the original Study Director is located. Additional
project staff are now located in both Newcastle and Brisbane and collaborators hail from all over Australia and some from overseas.

The organisational structure has evolved over time, so that the original project and management committee has been superseded by a Steering Committee, the DMG and the PSA. The role of each of these committees has evolved to manage efficiently the procedures and protocols of the ALSWH, discussed in further detail in Chojenta, Mooney & Warner-Smith (this issue).

**Conclusion**

In this paper we have described the types of skills and positions that might need to be considered for the successful operation of a longitudinal study. In addition to skills and qualifications, adaptability, flexibility and creativity are perhaps the most valuable traits that longitudinal study staff can have. The organisation needs to provide systems that enable both personal and organisational goal attainment in addition to ensuring succession planning and the preservation of institutional memory.

**Acknowledgements**

The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women's Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

**References**


**Renew Now For 2008**

**ISSN 1834-0806**  *International Journal of Multiple Research Approaches*  [www.ijmra.com](http://www.ijmra.com)

Volume 2/1-2 (2008) will publish two issues of approximately 96 pages in April and September.

Your renewed subscription will include collections of articles on two topics:

**Issues and Implications of Computer Assisted Multiple and Blended Research**

Edited by Eugenio De Gregorio (Rome) and Francesco Arcidiacono (Rome)

**Multiple Methods for Novice Researchers**

Edited by Elizabeth Waters (Melbourne), David Este (Calgary), Alicia O’Catthain (Sheffield) and Kathleen M. T. Collins (Arkansas)

To ensure uninterrupted delivery of *International Journal of Multiple Research Approaches*, please ensure that your librarian has renewed your subscription with a recognised subscription agent or directly with the publisher. Subscription details are available on the inside front cover of this issue.

---

eContent Management Pty Ltd, PO Box 1027, Maleny QLD 4552, Australia
Tel.: +61-7-5435-2900; Fax. +61-7-5435-2911; subscriptions@e-contentmanagement.com
www.e-contentmanagement.com
Accessing and disseminating longitudinal data: Protocols and policies

Catherine Chojenta, Rosemary Mooney and Penny Warner-Smith
Research Centre for Gender, Health and Ageing, The University of Newcastle, Callaghan, NSW, Australia

ABSTRACT
The development of clear rules and regulations around the access to and publication of data is imperative to the ongoing integrity of a longitudinal study. Careful planning of policies and protocols should be undertaken at the commencement of the study, and refined over time to incorporate the growing availability of data, and increasing numbers of external collaborators. In this paper we draw on the experience of developing data access protocols for the Australian Longitudinal Study on Women’s Health (ALSWH). We discuss the development of these policies and the organisational structure that manages them. We also discuss the record keeping practices implemented by the ALSWH and how the information stored can be used for both the review of study themes and also the production of research summaries and reports.

Keywords: Longitudinal studies; survey methods; research administration

Recruiting the right people and matching their skills to the work involved is critical to the success of a longitudinal study: these strategies were discussed in the previous paper in this issue (see Warner-Smith, Loxton & Brown). However, success is not dependent on personnel alone. Appropriate policies and protocols must be developed to underpin the effective analysis of collected data and the subsequent scientific outputs. Crucial questions, for example, are: Who will have access to the data? And how is access to be regulated?

Furthermore, any longitudinal study will have clear goals, generally established and refined in ongoing consultation with the funding body. As part of this process, it is likely that a number of themes, or topic areas for analysis, will be identified. The themes may evolve over time, as happened with the Australian Longitudinal Study on Women’s Health (ALSWH) in response to factors such as changing national priorities, the interests of the growing team of investigators, and comments and feedback from the study participants (see Byles et al. in this issue), but their identification acts as a lode star for the project. How can a study team guarantee that the work conducted by external researchers who are granted access to the data is congruent with the aims of the study? To ensure that data analysis is appropriate, and that themes or topics are covered, there must be effective policies in place to oversee data access and the publication of results, an organisational structure to manage these processes, and a record keeping system to monitor progress.
It is also important that the project interface with would-be collaborators is professional and accessible, given that the funding for longitudinal studies usually covers data collection but leaves little over for the analysis and dissemination of the data. Collaborators who wish to analyse the collected data and then to publish the findings are therefore very important to the productivity of a longitudinal study. Those who wish to undertake smaller studies of specific topics may also provide valuable extra coverage of study themes.

While we acknowledge that other longitudinal projects may have specific characteristics which call for different formal procedures, in this paper we draw on our own experience to discuss the development of appropriate policies and protocols. We consider firstly the management of access to ALSWH data, and secondly the overseeing of publications which use the Study data. The pro-forma documents and key organisational tools which have been developed are described, and we include flow charts illustrating the processes for establishing a new analysis and a 'substudy'.

**DEVELOPING POLICIES AND PROTOCOLS FOR DATA ACCESS, RECORD KEEPING, AND MONITORING OUTPUT IN THE ALSWH**

At the commencement of the ALSWH, the management of data analysis and publication was an internal process. For the most part, analyses were undertaken by the core researchers with the assistance of project staff. In this early phase, access was overseen by the Project Manager, and an informal record keeping system was established. It was not until the study team was approached by external collaborators and students over the years that it became apparent that a more formalised organisational system was necessary to manage and review the process of allowing access to data for analysis and the production of subsequent publications.

For example, although papers using project data have always been approved by project investigators before submission to journals, the formal committee established for this purpose was only set up eight years into the project, after this process was observed in another longitudinal study. Consequently, the system of keeping track of publications has changed a few times, resulting in the renumbering of individual papers on several occasions. This highlights the importance of considering and implementing these organisational systems at the start of your project. Table 1 describes some of the issues to consider when establishing a data access policy.

Further, the need to review the goals of the project to ensure that the collected data were being effectively analysed only became apparent some way down the track. It was then clear that this could not be achieved without a well-designed and rigourously maintained record keeping system. The maintenance of this information now allows both the ALSWH Steering Committee and the funding body to ensure that data analysis is focussed on key research themes, and that no critical areas are being overlooked.

Adequate record keeping strategies are imperative to the success of a longitudinal study. While the management of participant information is an obvious task in longitudinal data collection, keeping records of collaborating researchers and publication outputs is often overlooked. It can take some years from the initiation of an analysis to the publication of a paper on the findings. That paper may undergo a number of journal submissions, rejections, revisions, resubmissions, and so on. During that time it may 'drop off the radar' if there is no system to track its journey. In theory, another researcher could be granted access to the same data and unwittingly go on to reinvent the wheel, only to find when the first researcher's paper was finally published that the work had already been done.

Since the ALSWH began, over 420 people have been involved with the research in some capacity as investigators, collaborators, students or general staff. There have been over 175 analyses of project data, and 57 smaller projects ('substudies')
have been conducted with project participants. It has become increasingly important to keep track of who is analysing which variables in the datasets or investigating particular topics with the participants to prevent overlaps in research areas.

**DESIGNING A DATA ACCESS POLICY**

The ALSWH data access policy was created to manage and record new analyses being conducted by external collaborators. The process of granting permission to new collaborators was streamlined to better track collaborators and record information about the analysis. A searchable reference to all current and past analyses and substudies was also developed. The ALSWH data access policy is closely tied to our reporting requirements to our funding body, specifically the annual report and six monthly technical reports.

A key tool in the ALSWH is the Publications, Substudies and Analyses (PSA) Committee, which was established to manage the process of approval of new projects and papers. The PSA is a subcommittee of the ALSWH Steering Committee. Members of the Steering Committee are drawn from the team of investigators which includes people with a wide range of backgrounds and skills, including psychologists, sociologists, nutritionists, other allied health clinicians, medical practitioners, exercise physiologists, economists and so on. The external collaborators involved in the project represent an even broader disciplinary array. It is important therefore that the members of the PSA committee collectively span a variety of expertise in order to be able to comment on applications for access to the ALSWH data and to review academic papers prior to submission to a journal. The PSA chairperson reports regularly to the Steering Committee on the progress of new analyses and publications, brings to it any unresolved issues for discussion, and approves new applications for data access after considering comments from other committee members.

All collaborators are required to provide a six-monthly progress report for inclusion in the technical report. This requirement not only facilitates the creation of that report but also gives the PSA committee an opportunity to ensure that adequate progress is being made on all analyses. Although the occurrence is rare, approval can be revoked at this time if progress has been unsatisfactory.

The ALSWH categorise two levels of data access: analyses and substudies. An ‘analysis’ is any examination of ALSWH main survey data, and a ‘substudy’ is defined as a study of a subsample of ALSWH participants through an additional survey. For more information about substudies see Helman et al in this issue. Flowcharts illustrating the processes of application for data access, and the conduct of a substudy, are shown in Figures 1 and 2.

---

**Table 1. WHAT TO CONSIDER WHEN ESTABLISHING A DATA ACCESS POLICY**

- Who will manage the process of granting data access, and who makes the final decision? Do you need a committee or will a single chairperson suffice?
- How long is access granted, and what happens if a project is deemed ‘inactive’?
- How will you manage the potential overlap of research proposals? Your study team will need to consider how to mediate this potential hazard.
- Do you require the collaborator to feed information back to the research team at the conclusion of their analysis? Requirements might be for inclusion in annual reports, newsletters to participants or the project website.
- Will you charge external collaborators for data access and if so, how much?
- How much support is the research team able to give to external collaborators?
- Obtain legal advice from your institution’s solicitors when creating these agreements to ensure their legality.
- Consider the ethical requirements of your institution. Will each collaborator be required to obtain ethical clearance before proceeding with their analysis, or will the project’s existing approval be sufficient?
A new collaborator submits an Expression of Interest (EoI) form, detailing the proposed new analysis. If not already done, an ALSWH liaison person is nominated. This is usually a member of the ALSWH Steering Committee.

The EoI is given a sequential ID number and a standardised file name and circulated to the Publications, Substudies and Analysis (PSA) Committee.

Members of the PSA Committee review and comment on the EoI within two weeks.

The Chairperson of the PSA considers all comments from the committee members.

**Unresolved concerns, pending approval**

   - The PSA Chairperson will discuss the concerns with the collaborator

   - The PSA Chairperson will present the analysis at the next Steering Committee meeting for resolution

**If issues cannot be resolved by the collaborator, approval is not granted**

**No further concerns with EoI, approved**

   - The PSA Chairperson informs the collaborator of the decision to approve the analysis

   - The collaborator is provided with the necessary documentation:
     1. The Overview of ALSWH project and data
     2. Policy and procedure for data access, analysis and publication
     3. Memorandum of Understanding
     4. Privacy Protocol
     5. Data Request Form
     6. Qualitative Processing Protocols (if analysis uses qualitative data)

   - The collaborator signs and returns the necessary documents

   - The data is sent along with an invoice to the collaborator; at this time the cost for data is $100 however this is subsidised for student projects

   - The collaborator is required to supply six-monthly progress reports on their analysis for inclusion in the technical report (see Chojenta, Mooney & Warner-Smith in this issue for a description of these reports)

**If significant progress has not been made, it may be decided by the Steering Committee to revoke approval for the analysis or to reassign another collaborator to complete the analysis**

**If progress is satisfactory, collaboration continues**

**No concerns with EoI, approved**

**Figure 1. Flowchart for a new analysis**
The collaborator discusses the proposed substudy with an ALSWH liaison, and the Project Manager.

A preliminary meeting is held between the collaborator, the Project Manager and project staff who will be involved in running the substudy.

An Expression of Interest (EoI) is given a sequential ID number and a standardised file name and circulated to the Publications, Substudies and Analysis (PSA) Committee.

Members of the PSA Committee review and comment on the EoI within two weeks.

The Chairperson of the PSA considers all comments from the committee members.

**Figure 2. Flowchart for a new substudy**

Unresolved concerns, pending approval

- The PSA Chairperson will discuss the concerns with the collaborator.
- The PSA Chairperson will present the substudy at the next Steering Committee meeting for resolution.

No concerns with EoI, approved

- The PSA Chairperson informs the collaborator of the decision to approve the substudy.
- The collaborator is provided with the necessary documentation:
  1. The Overview of ALSWH project and data
  2. Policy and procedure for data access, analysis and publication
  3. Procedures for Substudies
  4. Memorandum of Understanding Privacy Protocol
  5. Data Request Form
  6. Qualitative Processing Protocols (if substudy uses qualitative data)

- The collaborator signs and returns the necessary documents.
- The collaborator must adhere to the approved substudy timeline. This process involves survey development, ethics approval, survey mailout, data collection and analysis.

Collaborators may be required to provide several forms of feedback:
  1. Prepare a short report for the ALSWH newsletter to participants on their research findings
  2. Six-monthly progress reports for inclusion in the technical report
  3. Detailed report for the annual report
  4. Feedback to substudy participants

- If significant progress has not been made, it may be decided by the Steering Committee to revoke approval for the substudy or to reassign another collaborator to complete the substudy.
- If progress is satisfactory, collaboration continues.

No further concerns with EoI, approved

- The PSA Chairperson considers the concerns with the collaborator.
- The PSA Chairperson will present the substudy at the next Steering Committee meeting for resolution.

If issues cannot be resolved by the collaborator, approval is not granted

- The collaborator signs and returns the necessary documents.
- The collaborator must adhere to the approved substudy timeline. This process involves survey development, ethics approval, survey mailout, data collection and analysis.

- Collaborators may be required to provide several forms of feedback:
  1. Prepare a short report for the ALSWH newsletter to participants on their research findings
  2. Six-monthly progress reports for inclusion in the technical report
  3. Detailed report for the annual report
  4. Feedback to substudy participants

- If significant progress has not been made, it may be decided by the Steering Committee to revoke approval for the substudy or to reassign another collaborator to complete the substudy.
- If progress is satisfactory, collaboration continues.

No further concerns with EoI, approved

- The PSA Chairperson considers the concerns with the collaborator.
- The PSA Chairperson will present the substudy at the next Steering Committee meeting for resolution.

If issues cannot be resolved by the collaborator, approval is not granted

- The collaborator signs and returns the necessary documents.
- The collaborator must adhere to the approved substudy timeline. This process involves survey development, ethics approval, survey mailout, data collection and analysis.

- Collaborators may be required to provide several forms of feedback:
  1. Prepare a short report for the ALSWH newsletter to participants on their research findings
  2. Six-monthly progress reports for inclusion in the technical report
  3. Detailed report for the annual report
  4. Feedback to substudy participants

- If significant progress has not been made, it may be decided by the Steering Committee to revoke approval for the substudy or to reassign another collaborator to complete the substudy.
- If progress is satisfactory, collaboration continues.

No further concerns with EoI, approved

- The PSA Chairperson considers the concerns with the collaborator.
- The PSA Chairperson will present the substudy at the next Steering Committee meeting for resolution.

If issues cannot be resolved by the collaborator, approval is not granted

- The collaborator signs and returns the necessary documents.
- The collaborator must adhere to the approved substudy timeline. This process involves survey development, ethics approval, survey mailout, data collection and analysis.

- Collaborators may be required to provide several forms of feedback:
  1. Prepare a short report for the ALSWH newsletter to participants on their research findings
  2. Six-monthly progress reports for inclusion in the technical report
  3. Detailed report for the annual report
  4. Feedback to substudy participants

- If significant progress has not been made, it may be decided by the Steering Committee to revoke approval for the substudy or to reassign another collaborator to complete the substudy.
- If progress is satisfactory, collaboration continues.

No further concerns with EoI, approved

- The PSA Chairperson considers the concerns with the collaborator.
- The PSA Chairperson will present the substudy at the next Steering Committee meeting for resolution.

If issues cannot be resolved by the collaborator, approval is not granted

- The collaborator signs and returns the necessary documents.
- The collaborator must adhere to the approved substudy timeline. This process involves survey development, ethics approval, survey mailout, data collection and analysis.

- Collaborators may be required to provide several forms of feedback:
  1. Prepare a short report for the ALSWH newsletter to participants on their research findings
  2. Six-monthly progress reports for inclusion in the technical report
  3. Detailed report for the annual report
  4. Feedback to substudy participants

- If significant progress has not been made, it may be decided by the Steering Committee to revoke approval for the substudy or to reassign another collaborator to complete the substudy.
- If progress is satisfactory, collaboration continues.
The guidelines and policies developed to cover the analysis and publication of ALSWH data are freely available on the project website, (www.alswh.org.au) and are also sent to new collaborators at the commencement of their association with the project. Not only do these policies clearly explain the expectations of collaboration and data access and publication, but also emphasise the privacy issues surrounding confidential participant information. Therefore all staff, students, and external data entry and mailing companies that have access to participant information are required to read and agree to the ALSWH data access policies.

In designing the ALSWH data access policy, the governing university research ethics committees were consulted to ensure that all ethical requirements were met by all policy documents. It is prudent for those embarking on a longitudinal study to consult the research ethics committee, and also the legal advisors at their institution/s. These experts should be consulted especially when devising the Memorandum of Understanding and Privacy Protocol: key documents that collaborators are required to sign. The literature relating to the creation of policy and procedures for data sharing in longitudinal research projects is limited however the following website may provide some helpful guidelines http://grants.nih.gov/grants/policy/data_sharing/.

**Proforma policy documents**

Table 2 describes the policy documents that have been created by the ALSWH research team.

The Expression of Interest (EoI) form which all intending collaborators must complete and submit for consideration if they want access to ALSWH data, or if they wish to conduct a substudy with a group of ALSWH participants, has undergone a lengthy process of development and refinement to ensure that all necessary information is provided by the collaborators, and that it is submitted in a standardised format. The current version of this document for an analysis of ALSWH data requires the following information:

- Name, title and institution of lead person/s
- Corresponding investigator and contact details
- Other investigators (including title and institution)
- Brief title of potential project
- If the research is the project of a student, student’s name, course and institution
- Identification numbers of previous ALSWH work conducted by the applicants
- Identification numbers of related ALSWH work, if known
- Age cohorts and surveys involved
- Brief outline of the topic, including a specific aim/research question or hypothesis
- Main variables or groups of variables to be analysed
- Expected outcomes and likely target audience (eg paper to be submitted to a psychology journal: paper to be submitted to National Rural Health conference)
- Time frame of the project (start/finish dates; expected paper submission dates)
- Source of funding for the work
- Other data sets required
- Names and email addresses of all people who will have access to the data
- Names of people who will provide the substantive expertise and input, statistical expertise and input, and qualitative analysis expertise if required.

The EoI form also includes a box for the use of office staff which records the number of the EoI; the date received; and the date on which the Memorandum of Understanding, the Privacy Protocol, and the Data Request Form were sent to the collaborators if the application was approved.

When the EoI is for a substudy, the following information is also required:

- Outline of substudy, including rationale, methods, and drafts of any surveys or interview schedules to be used.
• Ethical considerations, including participant confidentiality and care, institutions where clearance will be sought, and the status of existing ethics applications

• Budget
• Core ALSWH staff who will assist, and the tasks agreed upon.

### Table 2. ALSWH Data Access and Publication Policy Documents

<table>
<thead>
<tr>
<th>Document Name</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview of ALSWH project and data</td>
<td>• Details the methodology and describes the establishment of the project&lt;br&gt;• Contains important background information about the sample size, and dataset number of observations&lt;br&gt;• Is provided to all new investigators, collaborators, staff and students</td>
</tr>
<tr>
<td>Policy and procedures for data access, analysis and publication</td>
<td>• Describes the policies that have been developed around the publication of ALSWH data&lt;br&gt;• Describes the process for gaining approval for an analysis and then for the submission of a publication&lt;br&gt;• Is provided to all investigators, collaborators and students who will be analysing ALSWH data</td>
</tr>
<tr>
<td>Procedures for substudies</td>
<td>• Details the procedure for conducting a substudy using ALSWH participants&lt;br&gt;Qualitative processing protocols Details the conditions for using the ALSWH qualitative data which is collected on the back page of the survey under the question ‘have we missed anything?’&lt;br&gt;• Is provided to all collaborators who wish to analyse and publish qualitative data</td>
</tr>
<tr>
<td>Expression of Interest – analysis</td>
<td>• Is completed by all collaborators or investigators who wish to apply for permission to analyse ALSWH main study data&lt;br&gt;• Details the variables that will be used in the proposed analysis and the expected outcome of the analysis eg student thesis, journal article, conference presentation&lt;br&gt;• Sets out an expected time frame for the project, to enable the Steering Committee to ensure that the analysis will be completed and published in a timely manner&lt;br&gt;• Is sent to the PSA Committee for approval</td>
</tr>
<tr>
<td>Expression of Interest – substudy</td>
<td>• Is completed by all collaborators or investigators who wish to apply for permission to conduct a substudy using ALSWH participants&lt;br&gt;• Differs from Expression of Interest for analysis by asking about intended methodology, scale to be used, and sample</td>
</tr>
<tr>
<td>Data request form</td>
<td>• Is completed by the collaborator before data is posted to them&lt;br&gt;• Contains the information used by the data manager to ensure the correct datasets in the most useful format are sent to the collaborator</td>
</tr>
<tr>
<td>Memorandum of Understanding</td>
<td>• Is drawn up and signed by all parties once the Expression of Interest has been approved by the committee&lt;br&gt;• Details the project that will be undertaken and states the terms and conditions for collaboration, regarding ownership and use of the project data&lt;br&gt;• Forms a written agreement between the researchers and an ALSWH liaison person</td>
</tr>
<tr>
<td>Privacy protocol</td>
<td>• Details the rules for privacy of the participants’ information&lt;br&gt;• Is signed annually by all investigators, collaborators, staff and students before accessing ALSWH data&lt;br&gt;• Is also signed by any data collection or processing companies.</td>
</tr>
</tbody>
</table>
USEFUL TOOLS FOR DATA ACCESS AND RECORD KEEPING SYSTEMS

1 People and Projects Database

In order to manage the PSA approval process, a database called the People and Projects Database has been developed by ALSWH project staff. The database is used to produce reports for the Steering Committee, and also assists in the maintenance of the project website, and in the production of annual and technical reports (see Chojenta, Mooney & Warner-Smith in this issue). The database was created in Microsoft Access and cross-references each person who has access to the project data with each analysis and sub-study ever completed. Each project is also cross-referenced with the Publications Database (see below).

The ‘People’ data entry form is used to record key information about everyone who has access to ALSWH data, as well as all those who have had access at some time in the past, including investigators, collaborators, staff and students. Contact details are recorded, as well as staff positions and student enrolment details. The date that the person signed their last privacy protocol is also recorded in this database to manage the annual signing of privacy protocols. Each person in the database is given a unique identification number to facilitate the cross-referencing of ‘people’ data with ‘project’ data.

The ‘Projects’ data entry form is used to keep a record of both substudies and analyses, from the Expression of Interest stage, through the approval process, to completion. All researchers, past or present, who have been involved in the analysis are recorded on this form. Details of the datasets requested and those accessed are also recorded. Several other sub-forms can be accessed from this form, detailing the funding of projects as well as the details of substudies.

2 The project website

It is important that new collaborators are able to easily access information about existing analyses so they are able adequately to plan their application for data access, ensuring there is no overlap with existing analyses. Information about all ALSWH analyses is published on the project website.

The project website (for more detail see Chojenta, Mooney & Warner-Smith in this issue) is integral to the maintenance of communication with participants, stakeholders and researchers alike. It is important therefore to consider the amount of information to include on your website in regard to current projects and analyses. You should consider the ease of access to this information for new collaborators in order to ensure their proposal is unique, but at the same time, information on unpublished results should not be made publicly available.

Consider employing a databasing and website consultant to establish a searchable online database. Current technology allows information stored in Microsoft Access™ format to be summarised in html format. You may consider investigating other facilities available with your chosen website package, such as a MySQL™ platform database.

The ALSWH project database details information on project data access policies and protocols, for easy reference for new and existing collaborators. Lists of the titles, and researchers involved in all substudies and analyses can also be viewed in the public access section of the website.

3 Publications Database

In order to record the progress and status of publications and presentations relating to the project, ALSWH project staff have developed another Microsoft Access based database. The ALSWH Publications Database has a record of all papers from submission to acceptance, through to publication. Once a paper has undergone a PSA review (Table 3), and the collaborator has been granted clearance by ALSWH to submit the paper, a record of the publication is kept during each phase of its development. The information is entered into a data entry form, and several cus-
tomised reports have been developed so that information can be quickly and effectively summarised.

Papers saved in the Publications Database are cross-referenced to the People and Projects Database. This enables the tracking of research outputs from analyses and substudies.

**Management and Record Keeping of Publications which use ALSWH Data**

The publication of project findings in scholarly journals and at academic conferences is crucial to the ongoing academic quality of the project. In order to manage the timely publication of papers, the ALSWH has developed several protocols. As described in the *Procedures for Data Access, Analysis and Publication* document, all external collaborators are required to submit any publication to the PSA Committee for approval prior to submission to a journal. This process ensures that specific methodological details are described adequately, and that the data analysis was appropriate. A standard ALSWH acknowledgement has also been prepared for inclusion in all publications. This names the funding body and the research institutions, and thanks the participants for their contribution to the project.

**Process for approval of publications**

Collaborators must gain approval from the PSA committee before submitting any publications resulting from a substudy or analysis. As stated in the *Policy and Procedures for Data Access, Analysis and Publication*, PSA committee members will check the paper to ensure that the analysis appears correct and appropriate, that related work is adequately acknowledged, and that the paper does not overlap excessively with other papers published or in progress.

Copies of all versions of papers, from submission, to review, to publication must be submitted to the PSA Administration Officer for record keeping.

Both electronic and hard copies of all published papers are kept in the ALSWH research office for easy reference.

Factors to consider when developing a process for approving publications appear in Table 3.

**Thematic Review of a Longitudinal Study**

The recording of people, projects and publications information serves as a useful tool when creating summaries for the ALSWH website, and for annual and technical reports (for details about the methods used in creating these reports see Chojenta, Mooney & Warner-Smith in this issue). However, the record keeping databases are particularly valuable when undertaking a project thematic review, either internally by Investigators or in conjunction with the funding body.

The goals and themes of a longitudinal study should be made clear at the commencement of the study. However, as we have noted, over time it is likely that the research directions of a longitudinal study will change. It is therefore invaluable to be able to quickly and efficiently produce

---

**Table 3. What to consider when approving journal papers for publication**

- A paper may require two kinds of review before it is submitted for publication; a content review by an Investigator to ensure the methodology of the project is accurately represented, and with papers which include quantitative analysis a statistical review by a senior project statistician to ensure that the analysis is sound
- Consider who will oversee the approval process, and who will maintain the record keeping process
- Prepare a standard acknowledgement for inclusion in all publications. This will commonly include acknowledgement of your funding body, your institution and your participants. All investigators and external collaborators should be required to include this acknowledgement in all publications which use your data.
information on the scope of analyses and publications to ensure that current project themes are adequately covered, and to allocate resources to areas where there are gaps. For this reason, all new analyses and substudies are categorised by the existing ALSWH research themes when being entered into the People and Project Database. For further information about factors to consider when establishing project themes, see Table 4.

CONCLUSION

Our systems of record keeping in ALSWH have evolved over the past 11 years. The need for structured access and approval processes was not apparent to the ALSWH project team until we were well into the project. Consequently the system for tracking analyses and publications has changed several times, and processes will continue to be audited in the future to ensure they are still meeting our needs. This highlights the importance of considering and implementing these organisational systems at the start of your project. While this paper is not intended as an examination of the protocols employed by all longitudinal studies, it is hoped that the experiences shared here by the ALSWH will serve as a guide to others when establishing their own data access policies.

It is imperative when recruiting people to work on any new longitudinal study to ensure that adequate resources are allocated to the development and maintenance of record keeping structures to manage data access and the dissemination of findings.

The time and creative energy required to establish such structures is considerable, as is the painstaking attention to detail and perseverance to maintain up-to-date records and processes. The effort, however, is crucial to the quality and productivity of the study.

Acknowledgements

The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women’s Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

References

MySQL [Computer Software]. MySQL AB.

<table>
<thead>
<tr>
<th>TABLE 4. THEMATIC CONSIDERATIONS IN A LONGITUDINAL STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Develop clear themes in conjunction with your funding body</td>
</tr>
<tr>
<td>• Review these themes over time to ensure they still reflect the needs of the funding body, and the current research directions</td>
</tr>
<tr>
<td>• Categorise new analyses and substudies by theme, ensuring that any new project will help to fill in gaps</td>
</tr>
<tr>
<td>• Regularly review current projects by theme, to maintain coverage of core topic areas.</td>
</tr>
</tbody>
</table>

Accessing and disseminating longitudinal data: Protocols and policies

Volume 1, Issue 2, December 2007

INTERNATIONAL JOURNAL OF MULTIPLE RESEARCH APPROACHES

113
Longitudinal survey development and design

DEBORAH LOXTON AND ANNE YOUNG
Research Centre for Gender, Health and Ageing, The University of Newcastle, Callaghan, NSW, Australia

ABSTRACT
Many longitudinal studies collect data through self-report or administered surveys, either as the main source of data or as one of a set of data collection methods. Longitudinal studies offer special challenges for survey design including meeting diverse needs of investigators and stakeholders, developing consistent surveys that meet current and future needs, obtaining sensitive information in an ethical way, and producing a survey that is economically sound, easy to complete and has longitudinal integrity. This paper draws on the experiences of the Australian Longitudinal Study on Women’s Health (ALSWH) to provide some insight into the practical aspects of designing longitudinal surveys, including modes of administration, and the development of baseline and follow-up surveys. The ingredients for successfully conducting a longitudinal survey include extensive consultation, striving for balance between competing interests, review and documentation of all items and justification of new research questions. The commitment of an ever-evolving research team to these tenets contributes to the production of quality outputs which justify the ongoing contribution of the participants.

Keywords: longitudinal research; survey design; survey administration

Many longitudinal studies collect data through self-report or administered surveys, either as the main source of data such as the Australian Longitudinal Study on Women’s Health (ALSWH), or as one of a set of data collection methods. For example, studies that require the collection of biological samples may also collect some demographic and background information on risk factors or outcomes by survey data collection methods or medical record review. There are many texts available that describe techniques to assess the validity and reliability of survey questions and the psychometric properties of scales. For example, see Bowling (2001), McDowell (2006), and Streiner and Norman (1995). Similarly, advice on how to conduct cross-sectional surveys is generally available in undergraduate texts and courses. However, despite the growth in longitudinal studies worldwide, practical advice on how to develop surveys to support and facilitate longitudinal research is rarely reported. This paper, then, draws on the experiences of the ALSWH to provide some insight into the practical aspects of designing longitudinal surveys. Some of issues that are peculiar to longitudinal studies, such as integrity of data over time, maintaining consistency in design, and those tenets that contribute to a sound longitudinal survey instrument, will be the focus of this paper.
One of the first decisions faced by a research team is how to administer the survey and this will be discussed in the first section of this paper. Two phases of survey development will then be described: the initial phase which is concerned with designing the first (or baseline) survey; and the subsequent phases which involve designing the follow-up surveys. These phases form the second and third sections of this paper.

1. Mode of Administration

Surveys are commonly administered by mail, telephone or face to face interview. Over the past decade, with faster, more reliable and readily available Internet connections, online surveys are becoming increasingly popular. In a similar way, technology improvements have recently allowed hand held computers or Personal Digital Assistants (PDAs) to be used for survey data collection. Determining the relative advantages of these methods largely depends on who the participants are and what they are being asked to record. For example, mailed surveys may not be the most appropriate means of collecting information from people with known literacy problems. Similarly, people with lower socio-economic status might have poorer access to the Internet and so be under-represented in an online survey. Budgetary considerations will also play an important role in determining the mode of survey administration.

The ALSWH began its pilot studies in 1995 using self-complete mailed surveys. As the population to be surveyed included Australian women in three age cohorts ranging in age from 18 to 75 years, the method chosen needed to be acceptable to a wide range of women. Mailed surveys were chosen as the most efficient to meet the overall goal of the research, which was to determine the factors associated with healthy ageing among Australian women from a broad psychosocial perspective. Australia is the sixth largest country in the world in terms of land area, covering more than seven million square kilometres. As the study was designed to enrol a national random sample, with oversampling in rural and remote areas, the cost and logistics of telephone or face to face interviews precluded use of these as the main methods of data collection. Participants were randomly selected from the Medicare database, which at that time was the most complete enumeration of women in Australia (see Brown et al. 1998 and Lee et al. 2005 for more details). To enhance the participation of women with literacy problems, limited functional abilities (such as a visual impairment) and women who spoke languages other than English, telephone interviews were available upon request (for more information see Adamson, Young, & Byles in this issue). In 1996 online completion of surveys was not a viable option.

The ALSWH study team investigated the viability of using online surveys for the fourth survey of the Younger Cohort, which would be conducted in 2006. To determine how acceptable this would be to the participants, during the pilot test phase in 2005, participants were asked ‘Would you be willing to complete the survey online?’ Of the 120 women (then aged 27-32) who responded to the question, 44% replied that they would not complete the survey online. The main reasons given were that they did not have access to the Internet; they thought it might take too long; or they were concerned about confidentiality. Since mailed surveys would then have to be administered to women who were unable or unwilling to complete online surveys, there would be additional time and monetary costs. Therefore, the ALSWH has no immediate plans to administer surveys via the Internet although the options will continue to be monitored.

2. Baseline Survey Development

Designing a survey that will be in long term use involves creating a balance between competing demands. For example, increasing the length of the survey allows more questions to be included but increases participant burden. This is particularly important when trying to encourage completion of the first survey. The priorities of the research team and the funding body may differ;
there may be a discrepancy between quotes for the data collection process and the project budget; and a compromise between the desire for obtaining data quickly and the need to take the time to design a valid survey. With regard to these issues we will first discuss survey content, followed by format. We then discuss the process of pilot testing, which can be used to test the both the suitability of content and the efficacy of survey format.

2a. Developing baseline survey content
Research projects are often conducted within short time frames. While longitudinal studies might be undertaken over decades, gaining the first results from the baseline survey can be a pressing need. However the process of good survey development should outweigh the need for quick results. An inadequate development process can lead to inadvertent omissions or irrelevant or less than optimal questions being included. The baseline survey content can influence the response rate to the longitudinal study, as it is often the first contact between the research team and the participant. The face validity of the survey is very important. For example, some ALHW participants wondered why a survey about health should have questions about employment or level of education. The information brochure that accompanies the invitation to participate is very important to convey the reasons why certain types of questions are included in the survey.

From the outset, it is important to document the rationale for the inclusion of each item in the survey. The purpose is threefold. Firstly, if it proves difficult to justify inclusion, then it is likely that the item is not as necessary as it might have first appeared. Secondly, with the turnover of staff over time in a longitudinal study much of the ‘corporate memory’ to explain decisions might be lost. Longitudinal projects can continue for decades and it is unlikely that the team that starts the project will remain throughout the study period (Warner-Smith, Loxton & Brown in this issue). Thirdly, knowing why an item was included in the baseline survey can assist with evaluating its inclusion in subsequent surveys. As part of the documentation process it is also important to document scoring systems that are used with scales and measures, and any other pertinent information which might otherwise become lost over time.

The development of a baseline survey is important to establish sound measures of the factors that may predict the outcomes of interest as well as measures of those outcomes themselves. Well-defined research questions will assist with this process. However, with longitudinal research it is also important to consider the predictors and outcomes that might become important in the future. Consultations with funding bodies, government and relevant industry representatives and researchers, in conjunction with a thorough understanding of national and international trends in the broad research area to be addressed, can all contribute to the development of a sound baseline instrument.

In developing the ALWH baseline survey, consultations were held with people from lay, academic, and professional communities, and with policy makers to establish the areas that were most in need of research in women’s health. Among the consultants were clinicians whose specialities included physical and psychological health, social workers, women’s health, state and federal government health department staff, and researchers from many disciplines including epidemiology, psychology, sociology, medicine, gerontology, nursing, nutrition and biostatistics.

At the conclusion of these consultations, a draft 60-page survey was produced for each age cohort. As the length of the survey was prohibitive in terms of respondent burden, reducing the length involved a process of prioritisation of the items, based on their justification for inclusion. Where several measures of a factor had been proposed, the one considered best for the population under study was chosen, based on a review of the literature. A key feature in the decision making
process was assessing the items in terms of the original research themes in the study proposal (see Byles et al. in this issue), rather than exploring broader or more recently emerging research areas. This process involved a compromise from many members of the multidisciplinary research team and helped establish a spirit of cooperation amongst the investigators. These lessons are vital for the long-term success of a longitudinal study.

Focus groups with people drawn from the population of interest can also inform the content of surveys. Focus groups provide feedback on the acceptability of the items, important issues and concerns that may not have been covered, and areas where clarification is necessary, such as unfamiliar or confusing wording in the questions. Focus groups can also be used to gauge the level of interest and willingness of potential participants to take part in the project, and to offer insights into methods that can be used to encourage participation among their peers.

The ALSWH is a national study and therefore focus groups were held in rural and metropolitan areas with women who were within the age ranges that would be initially sampled (18-23 years, 40-45 years, 70-75 years). An example of the types of questions that were asked during the focus groups for younger women is provided in Figure 1. In this example, focus group participants talked about their feelings with regard to being asked about abortion. As a result of the feedback from focus groups, the baseline survey included a response option ‘I do not want to answer’ for some of the more sensitive questions. When asked how many times they had had a termination, only 0.5% of the Younger cohort chose not to answer the question. However, providing that option may have helped the women feel comfortable enough to continue to answer the remaining questions. The focus group results also indicated the importance of privacy and confidentiality to women in this age group. The confidential nature of the survey was stressed in all information sent to participants as a result of this finding. When draft questionnaires were available, these were also taken to focus groups for comment, so that feedback about wording, format, and content could be used to further inform the development of the survey.

The age of the people to be surveyed plays an important role in determining the number and types of questions that can be asked. For example, the ALSWH includes three cohorts that represent three distinct groups of women. The older cohort women were born from 1921 to 1926, the mid-age cohort were born from 1946 to 1951, and the younger cohort from 1973 to 1978. These women have grown up in very different time periods. The older cohort turned 18 during the Second World War, the mid-age cohort turned 18 in the mid-to-late 1960s, and the younger cohort turned 18 in the early 1990s. The three age cohorts have different sensitivities and different perspectives about what is and is not an acceptable question to be asked. For example, the younger and mid-age cohorts have been asked

---

**Facilitator:** Would you be offended if someone asked you if you had had an abortion? Or if you used contraception?

**Participant:** Well, not contraception, but I think abortion would make me touchier.

**Participant:** A lot of people see abortion as wrong and they can really judge you.

**Facilitator:** Would you answer the question on a questionnaire?

**Participant:** If it was anonymous.

**Participant:** I think it would be fair if there was an option not to answer it.

---

**Figure 1: Excerpt from a focus group of women aged from 18-23 years**
about their sexual orientation and whether they have had sexually transmitted infections. These questions may be unacceptable for many of the older women.

We aim to maintain consistency in the nature of the health and health-related questions we ask each cohort in order to facilitate comparison, but it is also important to focus on issues that are pertinent to the particular age cohort being surveyed. For example, in 2006 the younger cohort (then aged 28 to 33) were asked detailed questions about childbirth, while the mid-aged cohort have been asked about menopause in some detail. A further consideration is the effect of ageing during the course of the longitudinal study. The older ALSWH cohort was aged 70 to 75 at the start of the project and so their survey was produced in a larger font size and includes fewer items than the surveys for younger age cohorts. As this cohort has aged, it has been necessary to add to and modify the survey items. For instance, the baseline survey included few measures of disability. Over time a percentage of the cohort have developed functional limitations that could not be captured by the baseline measures. Further measures that better discriminate between greater levels of disability have needed to be added.

It is advisable to include validated survey instruments where possible, however there are a number of associated issues that should be considered. Firstly, validated instruments sometimes have copyright protection and their use requires approval and the payment of charges. If a decision is made to include these items, the charges will need to be included in the long term study budget as a recurring item. Similar measurement instruments that are freely available might be preferable where study budgets are limited or could become limited over the term of the study. Secondly, the majority of validated instruments have been tested in cross-sectional studies rather than over long time periods. The implications are that such measures might not be sensitive enough to detect changes over time. That is, a measure might detect differences at one point in time (cross-sectionally) but might not be sufficiently sensitive to detect differences over time that occur at an individual level. Of course, some traits that are measured longitudinally are found not to change and are stable traits. It is sometimes difficult to distinguish between stable traits and measurement tools that are not responsive to real change. Lastly, the time frames attached to the instruments that have been validated might not be appropriate for a longitudinal study.

Many survey instruments ask for responses to be considered in relation to a specific time frame, such as in the last four weeks, the last 12 months, or ever. These response time frames might need to be adjusted in longitudinal studies for the baseline and subsequent surveys, depending on the nature of the information required. Choosing the relevant time frame is reliant on the research questions. Is it important to know whether the situation is occurring now or recently (within a specified time frame), or is it more important to know whether the situation has ever occurred? The limitations of the time frames used are somewhat obvious, but still bear stating. An ‘ever’ time frame will not tell you how recently someone has experienced the item at hand. A limited time frame will not tell you if a person has experienced an event prior to that time.

In the baseline survey of a long-term study, historical (or retrospective) data may need to be collected. For example, with a cohort that starts at birth it may be useful to have family history and antenatal information. With an adult cohort, it may be important to collect information on medical history (eg the diagnosis of a chronic disease) and prior health risk factors (eg history of smoking). The manner in which these historical data are asked for and updated at later surveys has implications for the way in which the response can be categorised and analysed at later time points. Many of these issues relate directly to the time frames that are attached to questions within the survey. How will new cases of a chronic disease be determined when the longitudinal...
responses are collated? If a participant responds that they have ‘ever’ had a non-enduring condition (such as asthma), how do you know whether they have it now?

Where the relative importance of a time frame can not be determined, participants can be asked about a series of time frames. For example, the ALSWH asks participants about current and past use of tobacco. However, we have noted that including more than one time frame adds a layer of complexity to the survey than can easily lead to confusion and missing data.

In conclusion, the content of the baseline survey is vitally important and it is crucial to take sufficient time to develop that aspect of the survey. The format of the survey items has an equal role of importance once the content has been agreed.

2b. The format of the baseline survey

The format of the survey will be influenced by the population to be studied and the project budget. General guidelines of survey design apply to longitudinal studies (eg Brace 2004; Dillman 2000; Frazer & Lawley 2001; Presser et al. 2004). There are a number of issues that are specific to longitudinal studies, such as the importance of ‘branding’ and continuity in survey style. These issues have been addressed by the ALSWH and we discuss these in this section.

The ALSWH is planned to run for at least 20 years, with surveys of each cohort in a rolling three yearly cycle. The researchers therefore decided to ‘brand’ the project – that is, to create a logo and develop materials that would become familiar to participants and other stakeholders and be easily recognised over the life of the project. Logo and survey styles will stand the test of time provided that the colours and styles are not tied to a particular fashion or trend, making them less acceptable over time.

Although some aspects of the ALSWH survey format have changed over time, the basic layout exists today as it did for the baseline surveys in 1996. The front cover includes the study logo, the number and year of the survey and cohort (eg Fourth Survey of Younger Women 2006) and the coded participant identification number. Inside the front cover are the instructions for completing the survey. The inside back cover can be detached and consists of the consent form and a section for alternate contact details (see Adamson & Chojenta in this issue). The back of the survey is also a good place to add reminders such as ‘You may like to take a moment to check that you have not missed any questions or pages’. The back cover lists the contact details for the ALSWH. In this way participants become familiar with where to find information and how to complete the survey. Therefore, design of the baseline survey sets the tone for all future surveys.

Participating in a study that might last for several decades is a major commitment. Potential participants will not be encouraged to make this commitment if the first survey is difficult to complete or does not look professional. A primary consideration here is the length of the survey: the longer the survey, even if just due to greater spacing, the higher the level of commitment appears to potential participants. On the other hand, saving on length by squeezing more questions per page will make the survey appear cramped, unprofessional and onerous.

The ALSWH surveys have varied in length from 24 to 32 pages, although the 32 page surveys have prompted a number of participants to comment that the survey was ‘too long’. ALSWH survey items are well spaced, and the survey is divided into a number of sections with interesting headers. These headers were added following questions from pilot participants about what some questions had to do with health. For example, one section is headed, ‘Women’s health is about friends and family’ and contains questions about social support and marital status. About halfway through the survey, a watermark of a cup of tea/coffee with overlaid text advises participants that they are ‘half way through’. These considerations were designed to help participants feel as though they are advancing through the survey.
at an acceptable pace and could take a break when necessary (Dillman 2000).

The style and tone of the ALSWH surveys aim to be friendly yet professional. The professional appearance of any survey is helped by good quality printing and paper, however, this will be limited by cost and the mode of data capture. It is important to obtain advice from printers and companies being contracted to undertake the data capture, to ensure that their requirements are met. The mode of data entry may influence the number of survey items per page and the layout. Some scales and survey items will require specific layouts that will dictate the way in which they are included. The ALSWH survey responses were keyed in by hand at baseline but following a successful pilot test of 5,000 women in a large ALSWH substudy in 1997, the project moved to optical scanning of surveys from the first follow-up survey of the mid-age cohort in 1998. More details about data entry are available in Graves, Ball and Fraser (this issue).

Once the content and the format of a survey have been designed it is time to pilot test the survey.

2c. Pilot testing the baseline survey
Pilot testing of baseline surveys in longitudinal studies involves testing both the procedure and the survey instrument, and asking pilot participants to provide feedback on both aspects. In the ALSWH, pilot samples were drawn in the same manner as the main surveys except that they were recruited from two specific geographical areas. In all, five rounds of pilot testing took place from August 1995 until January 1996. However, the pilot studies were not all conducted in the same manner. For example, the first pilot study consisted of two groups, for each of the three age cohorts. One group from each cohort received a full-length survey, and the second group received a short survey, followed by a second short survey sent at a later date. This pilot study aimed to determine whether the length of the initial survey was an important factor in predicting response. Several other versions of the survey and accompanying materials were tested in subsequent pilot studies. The reminder protocols were also tested as part of this process. Did the response rate improve when the full package was mailed again or was a reminder card sufficient to elicit further responses? Response rates to each of the pilot tests were used to establish the manner in which the surveys would be administered for the main study.

The pilot participants were asked to complete a survey evaluation sheet that asked for feedback on various aspects of the survey. For example, pilot participants were asked to indicate questions in the survey that they thought were difficult to answer or too personal and they were asked to indicate their general level of satisfaction with the format of the questions. Comments from the participants in conjunction with an assessment of missing data, response patterns and response rates were used to decide the final format and content of the baseline surveys.

One important finding from the pilot surveys was the value of including a free response question at the end of the survey. The last question in ALSWH surveys asks, ‘Have we missed anything? If you have anything else you would like to tell us, please write on the lines below’. Although the ALSWH surveys are primarily quantitative in nature, over the years these free responses have provided a wealth of qualitative data that not only inform survey development but also provide valuable information about factors that affect women’s sense of health and well-being.

The ALSWH team found the pilot testing process to be invaluable. The pilot test groups have remained with the ALSWH, and have continued to evaluate the subsequent surveys produced by the ALSWH.

3. Development of follow-up surveys
Despite extensive planning of baseline surveys, unanticipated eventualities in longitudinal studies may lead to the need for changes in design, format or content in subsequent surveys. For exam-
ple, one facet of the ALSWH surveys that was not foreseen was the volume of material that would be returned to the project once the mailouts of the baseline surveys were undertaken. Dealing with over 40,000 returned surveys, the associated data entry and analyses across three age cohorts, in addition to answering an enormous number of enquiries, proved to be a heavy burden. Therefore, the ALSWH team decided that one cohort would be surveyed each year on a rolling cycle. This allowed time for the development of follow-up surveys to be undertaken, with only one cohort’s survey needing development each year. We now discuss the ongoing process of follow-up survey development.

3a. The process of survey development

Every year the ALSWH sends a follow-up survey to one age cohort, pilot tests the next survey of another cohort, and creates and cleans the dataset for the previous survey of the remaining cohort. For instance in 2007, the ALSWH team surveyed the Mid-age cohort for the fifth time, developed and pilot tested the fifth survey for the Older cohort, and prepared the dataset for the fourth survey of Younger women (which was collected in 2006). One question we are frequently asked is how we develop follow-up surveys. As with baseline surveys, it is important to consult with a wide range of people to ensure all relevant topics are covered and to pilot test the survey before sending it out to the main cohort.

The process of developing the ALSWH follow-up surveys is described in Figure 2. A call for comments on the content of the pilot survey is made to survey stakeholders including investigators, collaborators who have recently worked with the previous datasets, staff, representatives from government departments and other interested parties. Using the last follow-up survey as a starting point, email messages are sent seeking suggestions for new items to be considered for inclusion in the survey; items that could be deleted from the survey; and changes to existing items that might be warranted. The reasons why changes are made to the survey and issues that arise as a result of changes are discussed in the next section.

Note: The numbers 1 and 2 indicate the first and final drafts of the pilot survey respectively.

**Figure 2: Follow-up survey development process**
Qualitative comments that relate to survey questions from the previous survey of the cohort are reviewed. For example, the qualitative data from the Younger Cohort third survey (in 2003) indicated that polycystic ovarian syndrome was being mentioned by many participants, so the next survey included a question about the syndrome. Other preparatory steps include reviewing the levels of missing data in the previous survey, looking for items that were problematic following more complex analysis and comparison with other measures, and the responsiveness of items over time in the previous surveys.

Responses from the stakeholders, relevant qualitative comments and quantitative results were then collated into a single document in preparation for a planning meeting. Many items can be finalised during the planning meeting and for the residual items, working parties are established to conduct further investigation and provide recommendations within an agreed time frame.

The first draft of the survey is distributed for comments to the survey stakeholders, and then finalised shortly thereafter. The survey is formatted and proofed, and applications to the relevant Human Research Ethics Committees (HREC) are prepared and submitted. After HREC approval, the pilot surveys are printed and distributed. Results are usually available a few months later, and are reviewed by the same stakeholders who had input into the survey development process. The review and discussions are typically focussed on questions in the pilot survey with unacceptable amounts of missing data or poor discrimination. Poor discrimination occurs when one response is selected by the vast majority of participants, that is, it fails to distinguish between groups. For example, pilot survey questions that are found to be problematic can be reviewed and reworded before being retested in focus groups.

Final decisions for inclusions, deletions and changes for the next survey are made and the HREC applications are again submitted to cover any changes that were made to the survey as a consequence of the pilot survey process. Figure 2 then shows the overlap in ALSWH procedures: as the survey is printed and mailed out to one cohort, the pilot process for the next cohort commences.

All information about the survey development, especially additions, deletions and changes and the reasons for these are documented. This assists with HREC applications, but also creates a ‘paper trail’ for future data users to follow when analysing the survey data. These changes are incorporated into the data dictionary and the data map so that any changes in question wording, response options and format in each survey are noted.

It takes over a year to develop an ALSWH follow-up survey. The benefits of spending this time are difficult to assess but do result in good quality data, the prevention of formatting errors, being as inclusive as possible of diverse research areas and the ability to fully assess the value of including particular items. With that in mind, we now turn our attention to survey content in follow-up surveys.

3b Developing follow-up survey content.

Maintaining consistency in data that are collected across two decades or even longer is a challenging task. Times change, people change, and the political climate changes. Some research areas that were not important in 1996 are now very important in 2007. For example, the retirement plans of middle aged women, the health and well-being of single mothers, access to childcare and the use reproductive technologies have become more topical. By contrast, some survey items have become less important and it is difficult to justify continuing to collect some information. Similarly, changes to items might be warranted where they have not performed as well as expected. In this section we consider additions, deletions and changes to items in longitudinal surveys. All these actions have an impact on the integrity of the longitudinal data and the capacity to conduct longitudinal analysis and so should not be taken lightly.
As longitudinal studies progress, survey space appreciates in value. Sometimes there are compelling reasons to add new items to the survey. For example, with increasing attention being paid to retirement plans for women, requests to include more items about retirement have been made of the ALSWH. Additional items may also be proposed because the circumstances of participants change as they age. For example, participants might reach childbearing age, at which time it might be prudent to include more items about pregnancy and childbirth. Just as people age, so society and policy directions change. Keeping the research relevant to the current health and social policy context is important for the ALSWH, as many other studies. Additional survey items are suggested because social change has prompted inquiries from policy makers about what the study can provide. Often the wealth of longitudinal data that have already been collected, supplemented by some new questions or linked with other data sources, can shed light on research areas that a cross-sectional study cannot match. Research priorities can also change during the course of a longitudinal study, sometimes as a direct result of previous work conducted by the longitudinal study.

Each request for new items requires serious consideration. In order to add a question something else must be removed, at least for that survey. In considering the ‘worthiness’ of new items it is useful to consider the following points:

• Is the research question justified?
• Will the item contribute to knowledge in the field?
• How much space will the new items require?
• What existing items might need to be omitted to make way for new items and what are the consequences?
• Are resources available to analyse the new data?
• Could the research question be reasonably answered by a cross-sectional survey, rather than using the longitudinal resource?
• Are responses required from the entire cohort to answer the research question, or could the research question be answered with a substudy of participants (see Helman et al. in this issue for more information on substudies).

Under some circumstances the decision might be made to delete survey items. This can occur for several reasons. The items may have become redundant, for example, questions about childbirth will be irrelevant once participants are no longer of child-bearing age. Some items need only be asked once, such as country of birth (however we ask date of birth on each survey, as an identity check). Other items might be poorly answered or have a lot of missing data. For scales that have many questions, they may be included in some but not all follow-up surveys. Another reason for deletion occurs when there is little change in responses from participants over time. The following points can help in deciding whether an item is no longer required:

• Has the item produced responses with low rates of missing data?
• Has the item contributed to the research area?
• Is the item relevant to current policy and research areas?
• Has the item been responsive over time?
• If the item remains, will it contribute additional knowledge to the field?
• Are there sufficient resources available to analyse the data that are collected?

Revisions to the wording of questions often occur at the first follow-up survey where the time frame being referenced changes. For example, in the ALSWH baseline survey, women were asked to respond whether they had ever experienced a particular medical condition or event. In the second and subsequent follow-up surveys, the time frame for some items became the period since the last survey (three years). In this way, new occurrences of the relevant item could be measured.

If at all possible, revisions to existing items in a survey should be avoided or made in such a way that longitudinal integrity of data is maintained. That is, the revised item should have response
options that can be mapped in a consistent way to the original item to facilitate longitudinal analyses. For example, the ALSWH has measured relationship status (or marital status) in each survey of the Mid-aged cohort. As shown in Table 1, in the baseline survey women were asked if they were presently married, defacto (opposite sex), defacto (same sex), separated, divorced, widowed or never married. When the second survey was being developed, it was decided that relationship status should be consistent with how it was measured in the national Census taking place that year. Therefore, two items were included: one asked women about their formal registered marital status and another asked about their current living arrangements. These two items were then combined to create a variable that was compatible with data from the previous ALSWH survey, which could also be used to compare ALSWH respondents with the census. At the third survey, the item was returned to the original format which was easier to complete.

If it is not possible to retain longitudinal integrity, then revising an item can be viewed in the same way as deleting an item. The following points should be considered:

• Is the revised item an improvement on the original item?
• Will the revised item solve the problems that were apparent for the original item?

As a general rule, we often consider it better to retain an item in its original wording for the sake of longitudinal data. Even if the revised item is considered to be better in terms of wording, the loss of longitudinal data is a major disadvantage.

The process for making decisions about including, deleting or revising survey items will depend on the goals of the longitudinal study. The ALSWH process has been described, and this process is strongly linked to the priority research questions, the current policy climate, a consideration of emerging issues and a fundamental commitment to maintaining the longitudinal nature of the study and the data being collected. Longitudinal analysis cannot be done if the data items being collected do not allow for measures of change to be made.

**CONCLUSION**

Longitudinal studies offer special challenges for survey design, and the ALSWH is no exception. Among our challenges are meeting the needs of investigators, funding bodies and other stakeholders, developing surveys in an ongoing manner that are consistent with past surveys and that meet current and future needs, obtaining sensitive information in an ethical way, and

---

**TABLE 1: RELATIONSHIP STATUS MEASURES USED BY THE ALSWH ACROSS THREE SURVEYS OF THE MID-AGED COHORT: AN EXAMPLE OF ITEM CHANGES AND BACKWARDS COMPATIBILITY**

<table>
<thead>
<tr>
<th>Survey 1</th>
<th>Survey 2</th>
<th>Survey 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married</td>
<td>Married</td>
<td>Married</td>
</tr>
<tr>
<td>Separated</td>
<td>Separated</td>
<td>Separated</td>
</tr>
<tr>
<td>Divorced</td>
<td>Divorced</td>
<td>Divorced</td>
</tr>
<tr>
<td>Widowed</td>
<td>Widowed</td>
<td>Widowed</td>
</tr>
<tr>
<td>Never married</td>
<td>Never married</td>
<td>Never married</td>
</tr>
<tr>
<td>Defacto (opposite) AND</td>
<td>AND</td>
<td>Defacto</td>
</tr>
<tr>
<td>Defacto (same)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living in reg’d marriage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living in defacto relationship</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not married</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
finally producing a survey that is economically sound, easy to complete, straightforward to process and that has longitudinal integrity. The ingredients for successfully facing and succeeding in these challenges include an extensive consultation and communication process, striving for balance between competing interests, detailed review and documentation of past and current items and justification of new research questions. Underlying all of these activities, there must be a constant commitment to protecting the integrity of the longitudinal data by making only minimal changes to the items. The strong commitment of the ever-evolving research team and the allocation of responsibility for the survey content and the subsequent analysis and write-up of the results, contribute to the longitudinal study being able to produce high quality outputs and hence justify the ongoing contribution of the participants.

Acknowledgements
The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women’s Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

References

Brace I (2004) Questionnaire design: how to plan, structure, and write survey material for effective market research, London: Kogan Page
Recruiting for a longitudinal study: Who to choose, how to choose and how to enhance participation?

LYN ADAMSON, ANNE YOUNG AND JULIE ELLEN BYLES
The Research Centre for Gender, Health and Ageing, The University of Newcastle, University Drive, Callaghan NSW, Australia

ABSTRACT
There are many methods for establishing and recruiting participants for longitudinal studies. Mostly, the participants will be sampled from a population, and the study will need some list or methods for identifying and selecting the people to be invited to take part in the study. The choice of methods for selecting and recruiting participants will depend on the nature of the research question and the data to be gathered and on practical considerations such as cost and feasibility. In this paper we consider various sampling frames and methods, and we describe in detail the strategies used to contact women from the Australian Longitudinal Study on Women's Health (ALSWH) sample and to encourage their participation in the study.

Keywords: sample; sampling frame; recruitment; longitudinal; cohort; selection

This paper outlines the strategies and methods used to enhance recruitment for a longitudinal study. One of the advantages of the prospective nature of a longitudinal study is that the study factors can be defined as the investigator wishes and information can be collected on known confounding variables. Longitudinal studies help provide information about the causation of disease by allowing the temporal sequence between exposure and disease to be established. They also provide the most direct measurement of the risk of developing disease and a range of potential risk factors can be measured over time. The longitudinal nature of the study allows study factors to change over time in line with emergent issues. Such studies also have the advantage of allowing particular groups of interest to be recruited at baseline in sufficient numbers to allow meaningful analysis of results over time. The challenge in commencing a long-term study is to define the population of interest, examine the potential sampling frames and methods to recruit participants and consider ways to enhance participation and retention in the study.

WHO TO CHOOSE?
One of the first steps in establishing a longitudinal study is to identify the cohort of people who will be followed in order to answer the research questions of interest. Participants for the longitudinal study may be selected from a variety of sources and the choice will depend on a number of scientific and logistical considerations. These include the prevalence among certain population groups of the exposures of interest, the need to obtain complete and accurate information on
exposure to risk factors, and the likely success of being able to follow-up particular study participants over long periods of time.

For risk factors that only occur rarely (such as those related to particular industries, occupations, or environmental factors), it may be more efficient to choose a group who are exposed to the potential causative agent at a much higher level than the rest of the community and then recruit other participants from the general population for comparisons.

The Australian Longitudinal Study on Women’s Health (ALSWH) was designed to address broad questions about the factors that influence women’s health and well being. This required that the study included a broad cross-section of women in Australia, including women from different socio-economic backgrounds, with different occupations, and from different parts of the country. As a particular focus of the study was the health of women living in rural and remote areas, a method to select women according to area of residence was important.

A further consideration in selecting the cohort is whether to have a single cohort, or a series of cohorts. Many cohort studies have only a single cohort. An example is the 1946 British Birth Cohort Study (2007) which includes only people born in that year. In contrast, the ALSWH includes three five-year age cohorts as this allows exploration of differences in the experiences of three generations of women. Initially, only cross-sectional differences could be explored, but as the women age and the Younger cohort approaches the initial age of the Mid-aged cohort, and as the Mid-aged cohort approach the initial age of the Older cohort, then both age and cohort effects can start to be disentangled.

**HOW TO CHOOSE THE SAMPLE**

Participants for a study can be recruited by enlisting volunteers or by probabilistic sampling and invitation. In order to select a random sample of the population of interest, a complete sampling frame, or list of all eligible people is required. In many circumstances it is not possible to have a complete enumeration of the population, so other sources of eligible people are used. These may include voting lists (which may be out of date and not include recently arrived residents and people not eligible to vote for other reasons), medical insurance lists (which may exclude some members of the population of interest, particularly if related to income or age), telephone lists, and so on. Where the study participants have particular characteristics of interest, such as age, sex or medical condition, then information on those variables either should be available before sampling or else people are sampled and then checked for eligibility. A common method of recruiting people for longitudinal studies is to ask for volunteers that satisfy particular selection criteria, such as being in the correct age range and/or living in a particular area. The disadvantage of using volunteers rather than a random sample is that the conclusions of the study may not be generalisable to the whole population where the study sample was not representative, and biases can not be quantified.

Most studies follow either a whole population or group, or a random sample of a larger population. Following a random sample allows the inferences derived from the cohort to be applied to the reference population. The choice of sampling frame will depend on the nature of the study and the questions of interest. The original cohort of the Nurses Health Study (2007), for example, included 121,700 married female registered nurses aged 30–55 years and living in 11 large American states. While the results are able to be applied to women in general, not just nurses, there are some limitations in terms of the range of socio-economic and marital status. Similarly, the British Doctors Study invited all male physicians registered in the United Kingdom in 1951 to participate. Again the results may be generalisable to men who are not physicians, but the homogeneity of the sample in terms of occupation and social status means that little information can be derived about the influence of these factors on health, as there is no comparison group with other levels of these fac-
A number of sampling frames were available to the ALSWH, including the potential to recruit through one of the major women’s magazines. This option was not used as it would not have provided a population sample with known and measurable biases. The other main alternatives were the electoral roll (or voting register) and the enrolment database of Medicare (the national health insurance system). The choice of the Medicare database was determined by the fact that this frame should have been the most complete, since it technically included all people in Australia, including non-residents, regardless of age or socio-economic status (Lee et al. 2005). The other advantages were that it provided the technical capacity to link to Medicare health service utilisation data, and could provide aggregate health care utilisation information for groups of women who did and did not reply to the invitation to participate in the study. This information was useful in ascertaining the initial participation bias by providing answers to questions such as: did the women who accepted the invitation to join the study have the same patterns of health service utilisation as women who did not participate? Other studies such as the New South Wales 45 and Up Study (2007), and the Longitudinal Study of Children (2007) have also used the Medicare data base as the sampling frame.

The electoral roll has been used in many Australian longitudinal studies. The roll includes all adults who are eligible to vote in federal and state elections, that is, any person who 18 years of age or over, and who is an Australian citizen, or was a British subject on a Commonwealth electoral roll as at 25 January 1984. According to recent assessments of electoral roll accuracy, over 90% of eligible voters are recorded at their correct address (Australian Electoral Commission 2005). The Canberra Longitudinal Study of Ageing (2007) have also used the Medicare data base as the sampling frame.

Other studies are interested in specific medical conditions. For example, the Australian Diabetes, Obesity and Lifestyle (AusDiab) Study (2007) which commenced in 1999-2000 is providing benchmark national data on the prevalence of diabetes, obesity, hypertension and kidney disease in Australia. AusDiab is a field survey involving both physical testing as well as questionnaires. A team of researchers went to each of the 42 randomly selected urban and rural
areas (Census collector districts) around Australia to individually test each of the 11,247 individuals who participated in the baseline study.

Table 1 provides an overview of the sampling techniques in a selection of longitudinal studies. These studies range from recruiting babies and children via random population sampling to seeking volunteers from particular communities to be followed over time. Each study has a different objective and the study design and subsequent analyses will be specific to that objective. A thorough review of the sampling methods used in a number of Australian longitudinal studies is provided in a report produced by the Australian Institute of Health and Welfare (AIHW) (Logie, Hogan & Peut 2004).

Having established the sampling frame, the next step is to estimate the number of people needed to ensure that the longitudinal study will have the statistical power to make valid conclusions. It is not ethical to begin a study that has poor prospects of being able to scientifically test the hypotheses of interest. Pilot studies can help to estimate the likely response rate to invitations to participate and also test the willingness of participants to commit long-term to the study. Pilot studies are also useful for estimating the variability in the measures of interest, which is an important consideration in sample size calculations. Further information about the factors that affect sample size and the resultant power of a longitudinal study is presented in Young, Powers & Wheway in this issue.

### Table 1: Examples of Sampling Frames and Methods to Recruit Participants in Longitudinal Studies

<table>
<thead>
<tr>
<th>Name of study</th>
<th>Population of interest</th>
<th>Sampling frame and method of recruitment</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian Longitudinal Study of Ageing, Bedford Park, Australia</td>
<td>People aged 70+ in the Adelaide Statistical Division</td>
<td>Randomly selected from within the Adelaide Statistical Division. Both community and institutionalised individuals were included.</td>
<td>2,087 people</td>
</tr>
<tr>
<td>45 and Up, Sydney, Australia</td>
<td>People aged 45+ in New South Wales state</td>
<td>Randomly selected from the Medicare (national health insurance) database, with oversampling of those born overseas and of rural residents</td>
<td>More than 250,000 people</td>
</tr>
<tr>
<td>The Dubbo Study of the Health of the Elderly, Dubbo, Australia</td>
<td>Non-institutionalised residents born before 1 Jan 1930 (then aged 60+) first interviewed in 1988</td>
<td>Dubbo local government area eligible residents</td>
<td>2,085 Dubbo residents (1,233 men and 1,572 women)</td>
</tr>
<tr>
<td>Florey Adelaide Male Ageing Study, Adelaide, Australia</td>
<td>Men aged 35–80 in Adelaide’s north-west suburbs</td>
<td>Random sampling using electronic residential telephone directory</td>
<td>1,000 men</td>
</tr>
<tr>
<td>Australian Longitudinal Study on Women’s Health</td>
<td>Australian women in 3 age cohorts (18–23, 45–50 and 70–75)</td>
<td>Randomly selected from the Medicare (national health insurance) database. Oversampling in rural and remote areas</td>
<td>Almost 40,000 women</td>
</tr>
<tr>
<td>Longitudinal Study of Australian Children (Growing Up in Australia)</td>
<td>Infants aged 4–13 months and children aged 4 years 4 months–5 years 1 month at selection</td>
<td>Medicare (national health insurance) database</td>
<td>5,000 infants and 5,000 children</td>
</tr>
</tbody>
</table>
THE PRACTICALITIES OF RECRUITING FOR A LONGITUDINAL STUDY

This section outlines the recruitment strategies and methods which were used in the ALSWH. Collecting longitudinal data by mailed survey from large populations requires extensive planning and research prior to the initial recruitment stage to ensure maximum response rates. All the strategies described here were pilot tested on representative groups of women before being applied to the recruitment of women for the twenty year long project.

Due to privacy and confidentiality concerns many researchers are unable to invite potential participants to take part in research directly, and may need to rely on the holders of large databases to invite the participants on behalf of the research team. Even where researchers have direct access to names and addresses, potential participants are unlikely to know the researchers, the research process or what the process will ask of them. This situation raises a number of communication issues. When recruiting anonymous participants the project is being marketed to an unknown audience. It is therefore important that the initial message is easily understood, credible and provides some incentive (not necessarily a material reward), for a participant to be willing to consider undertaking the commitment to a longitudinal project. Understanding may be promoted by delivering clear simple messages. Use simple phrases, short sentences and avoid jargon (NHMRC 2000) when developing recruitment materials. Enhance the credibility of the project by demonstrating strong links between the research team and known reputable organisations, such as universities or government departments. Use non-material incentives by emphasising the contribution one person can make to change the bigger picture, such as informing health policies or services.

DATABASE ACCESS

Australia’s nationalised health care system is known as Medicare Australia. This centralised system maintains records of health care use by Australian citizens and registers of the providers of medical services. The Medicare database was chosen as it is the most complete record of people living in Australia. The fact that the sample frame provided opportunities for linking the participating women’s survey data to their health service use was a further reason for the selection of this database. The database was used to randomly select women in three specific age groups, to receive an invitation to participate in Women’s Health Australia.

THE INVITATION

Medicare Australia (previously known as the Health Insurance Commission) sent the letter of invitation to women on behalf of the research team. The identity of the selected women remained unknown to the researchers until a signed consent form or completed survey was returned to the research team. This method of recruitment necessitated development of strategies and protocols to ensure maximum encouragement to participate was made at the initial point of contact and at any subsequent contacts with the invited women. While this recruitment strategy protects the privacy of individuals, it limits opportunities the research team has to encourage a response to the mailed communication, unless the participant contacts the research team (eg by telephone).

MAILING PROTOCOL (AN APPLICATION OF THE TOTAL DESIGN METHOD BY DILLMAN 1978)

The mailing protocol for the ALSWH was developed after extensive pilot testing of a number of methods, and was based on the mailing strategies of Dillman (1978). The importance of planning cannot be stressed strongly enough. If the selected recruitment database allows access by the research team to names and addresses prior to requesting consent, strategies and protocols for encouraging participation (eg by telephone or
email contact) may be different from the ones used in the ALSWH.

Dillman’s recruitment method was devised for American postal conditions. As postal services and costings differ across countries a modified strategy to suit Australian postal conditions was developed. These modifications related to the timing of reminders and the type of mailing service used. All packages and reminders were sent by normal bulk mail service delivery, and mailouts were staggered across several months. Each of the three age cohorts was mailed at separate times. This allowed the research team to respond to each age cohort appropriately, and to ensure targeted media strategies were focused to each specific age group.

THE FIRST ALSWH SURVEY
A four stage mailing protocol was used for the baseline surveys of the ALSWH. Details of the contents of each package are described below.

MAILOUT ONE
This comprised six items as follows:

• Outer window faced envelope with the logo and return address of Medicare Australia.

  The package was mailed on behalf of the research team by Medicare Australia. The Medicare logo and return address appeared in black and white on the outer envelope. Packages not delivered were then returned by Australia Post to Medicare Australia, ensuring the anonymity of all invitees.

• A non-personalised invitation letter from the director of the project on project letterhead.

  Particular attention was paid to the order of the items in the package and the wording of the introductory letter (Dillman 1978). The introductory letter was the first item seen when the package was opened. As first impressions are important, the wording and presentation were crucial factors in maximising the response rate. As the research team were unaware of the names and addresses of invitees these letters were not personally addressed. However letters were printed on the letterhead of the project, Women’s Health Australia, and signed by the project director. The wording and font size of the letters for the mid-age and older age cohorts were the same and focused on the aspects of the project that might be of particular interest to women in these age groups, such as their interactions with the health care system and the benefits their participation might have to generations of Australian women. To encourage young participants to take part, the wording was changed slightly, with emphasis on the comprehensive nature of other aspects of the project as it encompassed not only physical health but also lifestyle and experiences with the health care system. Initial pilot testing with focus groups helped to inform the content of this letter (see Loxton & Young in this issue).

  As the scale of the project and ethical considerations prohibited the offering of material incentives to participants, it was hoped that the focus on opportunities for individuals to make a difference to future health policy would be an incentive to take part (Dillman 1978).

• An information brochure

  The information brochure was printed on coloured A4 paper, folded to DL size. The distinctive green and purple logo of Women’s Health Australia appeared on the front cover to ensure the invitee was instantly aware of the ‘branding’ of the project. This logo has subsequently appeared on all documentation sent to participants (see Chojenta, Byles, Loxton & Mooney in this issue). The brochure took the form of questions and answers covering all the potential concerns the invitee may have, based on the issues raised in the pilot focus groups and surveys. The back of the brochure outlined the project in nine languages and offered the Freecall number to obtain further information.
A personal consent form
A personally identified consent form on an A4 sheet asked the invitee to complete and return it, confirming their name, current address and phone number. It included an invitation to call the Freecall number to complete the survey by telephone in English or in another language if the invitee was unable to complete the written survey in English; a reminder to return the consent form with the survey in the reply paid envelope provided; and an option to withdraw her name from the mailing list. There was also a contact name, address and phone number for the university human ethics officer overseeing the project, in case the invitee had questions, comments or concerns she did not wish to raise with the researchers.

A copy of the questionnaire
The questionnaire was printed on coloured paper with black font. The logo appeared in black on the front cover with a brief message maximising the importance of each individual’s contribution – ‘We want to hear from YOU, YOUR views and experiences are important.’ The back cover thanked the participant and reminded her to include the consent form, as without this information the participant would remain anonymous to the research team. The format and colours for all subsequent surveys have been changed to allow for different forms of data entry (eg scanning) and to further encourage participation by standardising the image of the project format.

Reply paid envelope with a return address to the research team
The reply paid envelope had the logo and address of the project in the top left hand corner with the name ‘Women’s Health Australia’ included in the reply paid address. Although the completed surveys were returned to a data collection agency, the placement and emphasis on the project name and logo reaffirmed to the invitees that the levels of confidentiality outlined in the invitations would be maintained.

Mailout Two
This was sent one week after Mailout One. It included:

Thank you/reminder card
The ‘thank you/reminder card’, which thanked participants for completing their survey and prompted non-respondents to do so, was mailed seven days after the initial mailout. This procedure has since been altered. ALSWH no longer mails a reminder seven days after the initial mailout, but varies the follow-up across age groups, with a wait of at least 10 to 14 days for the older cohort before sending this reminder. This procedure was modified in response to calls to our Freecall 1800 number by members of the older cohort who were confused by the speed of receipt of the card and were stressed if they had not managed to complete the survey. The project logo appears in the centre top of a DL size coloured card, with slightly varied wording for each cohort. Each card carries a thank you for those who have completed the survey and returned it, and an encouraging statement outlining the importance of the project for those invitees who may have been hesitant to complete it. This is in keeping with Dillman’s (1978) Total Design Method.

Mailout Three
This was sent four weeks after Mailout One. It included:

Total replacement package to non-respondents
Four weeks after Mailout One, invitees who had not responded to either mailout one or to the thank you reminder card, received another complete package with a letter from the Project Director that thanked those who had already completed the survey and reminded those who may be hesitating of the importance of the project and the value their individual contribution could have on improving health services for Australian women. A DL size piece of coloured paper with black font emphasised that this was a replacement package.
**Mailout Four**
- **Reminder card**

Four weeks after Mailout Three, all participants who had not responded in any way to the previous three mailouts received a reminder card. This card was DL size on coloured paper with black font. It stressed the value of the program, the need for ‘women from all walks of life’ to complete it and emphasised that there was still time to participate. Many invitees who had made a decision not to respond but had not communicated this to the project were annoyed by the insistence of this final approach. Invitees who phoned indicated that they found this last card irritating. Although this procedure follows Dillman’s advice, making this last mailed approach should be considered carefully as, in the case of ALSWH, the cost of printing, postage and telephone charges clearly outweighed the potential benefits. The number of invitees who chose to take part as a result of the final mailout was minimal, and the major irritation it caused was clearly counterproductive.

**Media Campaign in Support of the Forthcoming Mailouts**

At baseline a concentrated media campaign was devised to coincide with the initial mailout for each cohort. As the cost of a national advertising campaign was not included in the project funding, strategies for achieving maximum media coverage with minimal cost were developed.

Press releases outlining the project were faxed to all media outlets across Australia, specifically targeting regional and rural areas. As Australia has a large number of ‘free’ newspapers distributed in local communities, good editorial coverage was achieved nationally.

Paid advertising costs were prohibitive, so advertisements announcing the survey (with encouragement to participate) were placed in publications that allow advertisements at minimal or no cost as a community service. Publications such as the Seniors Weekly and the free ‘street press’ papers targeted at younger people were used for advertising just prior to the first mailouts for the older and younger cohorts.

Feature articles based on press releases from the results from the pilot surveys were included in periodicals that were distributed to general medical practitioners. The press releases achieved good responses from syndicated commercial radio, television and the government broadcaster (the ABC). Several of the chief investigators from the research team were interviewed. This campaign was particularly effective on local and rural radio.

The younger cohort required additional media exposure to encourage response. University students undertaking a degree in communications that had a work experience component worked with the researchers to create a short radio documentary that highlighted results from our pilot studies. The students also created several commercial style radio advertisements. This documentary and the advertisements were played on a national community radio network. In addition the national broadcaster’s flagship station aimed at the younger age group (Triple J) made a short radio documentary presented by a well-known journalist.

If the research project is based within a university there may be courses that require work experience or internships within a course work degree to help design recruitment strategies at minimal cost. Examples of appropriate courses may be public relations, communications, graphic design, business studies and journalism.

**Responding to Enquiries from Invitees**

The telephone provides an opportunity to answer queries, cement relationships and gain additional information such as changes of address, email addresses or contact details for family and friends should the participant move. At the time of recruitment however, the telephone was used to encourage participation and to correct addresses,
as well as to take details from invitees who did not want to participate.

A Freecall number was provided for invitees who required additional information or wished to decline the invitation to participate. This call incurred no cost to the invitee. Each call was answered by a trained research assistant who recorded each call on a paper log. Names of those invitees who rang to decline participation were recorded to ensure they did not receive any further packages or reminders. Their main reason for refusing to participate was also recorded.

Information collected from the first few hundred callers was used to develop flip-charts that helped the telephone operators to offer standardised responses to future callers. The flip-charts included responses to frequently asked questions (see Table 2), contact phone numbers for Medicare Australia who had provided the initial names and addresses, and for other women’s health services, and some sample positive statements to encourage participation (see Table 3).

The telephone was also used to encourage invitees who could not complete the survey in English to complete the survey in their own language. As participants in the ALSWH speak over 80 different languages, this was a considerable challenge. We could not afford to have interpreters for every language instantly available to respond to each enquiry, but by being aware of the most common languages spoken, and being able to re-call them quickly to clarify issues, we were able to recruit a large number of culturally and linguistically diverse participants.

The majority of calls taken were to decline the invitation to participate. The Freecall number was critical for gathering information about the reasons women gave for not wishing to take part in the study. The reasons were noted and coded into major categories for statistical purposes (see Table 4). Reasons for non-participation could not be collected unless the women chose to contact the research team, as the researchers were blind to the names and addresses of the selected women.

### Table 2: Examples of Frequently Asked Questions

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why are you doing this?</td>
</tr>
<tr>
<td>Why not men’s health?</td>
</tr>
<tr>
<td>What do I have to do?</td>
</tr>
<tr>
<td>Who is providing the funding?</td>
</tr>
<tr>
<td>Where did you get my name?</td>
</tr>
<tr>
<td>Will all the information I give be confidential?</td>
</tr>
<tr>
<td>Why does it have to run for so long?</td>
</tr>
<tr>
<td>Do I have to take part?</td>
</tr>
</tbody>
</table>

### Table 3: Affirmative Statements

<table>
<thead>
<tr>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘This research is unique it is the first national project to research women across the lifespan’</td>
</tr>
<tr>
<td>‘You will be doing something to benefit your grandchildren’</td>
</tr>
<tr>
<td>‘This is a wonderful opportunity to share your health experiences’</td>
</tr>
<tr>
<td>‘The views of young women are the key to clarifying how today’s changing lifestyles may impact on health in the longer term’</td>
</tr>
<tr>
<td>‘You will be informing health policy, this is a chance to contribute’</td>
</tr>
</tbody>
</table>

### Table 4: Reasons for Declining the Invitation

<table>
<thead>
<tr>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Too busy</td>
</tr>
<tr>
<td>Survey is an invasion of privacy</td>
</tr>
<tr>
<td>Data linkage issues</td>
</tr>
<tr>
<td>Questions are too personal</td>
</tr>
<tr>
<td>Dislike of longitudinal study design</td>
</tr>
<tr>
<td>Survey too long and difficult</td>
</tr>
<tr>
<td>Invitee is too healthy</td>
</tr>
<tr>
<td>Invitee is too old</td>
</tr>
<tr>
<td>Invitee is caring for another person, no time</td>
</tr>
<tr>
<td>Nomadic lifestyle</td>
</tr>
<tr>
<td>Requests payment for completing survey</td>
</tr>
</tbody>
</table>

The remaining calls fell into several categories, including reassurance for invitees who were unsure about responding, changing incorrect address details or requesting a replacement for a lost survey. Every opportunity was taken to encourage invitees who rang to participate, without pressuring the invitee. Affirmative statements were included in the flip-charts provided to the telephone operators.
Many phone calls were taken from the ‘maybes’. These invitees were targeted and responses prepared for a wide variety of possible scenarios. We were frequently asked if other family members or friends could complete a survey and take the place of the original invitee or if non-invitees could complete the survey on behalf of family members who may have had disabilities that prevented them from completing the survey alone. There was a need for consistency in responses to these situations.

It is vital to answer the calls to the Freecall number promptly. Obtain as many inward lines as is practical. Piloting all recruitment strategies and keeping records will help assess how many inward lines needed to ensure all calls are answered promptly. An indecisive invitee will become a non-participant if they are forced to keep calling a number, especially if another reminder or package arrives in the meantime.

**Conclusion**

Maximising response rates at the initial recruitment is vital to the success of longitudinal studies. The lessons learnt from the recruitment of the main cohort of the ALSWH were used as a basis for encouraging retention in the project and were instrumental in creating successful procedures and protocols for the recruitment of participants into subsequent substudies (see Helman et al. in this volume). Importantly, recruitment of the women into the study involved more than mailing thousands of letters and questionnaires. First, a number of pivotal design issues had to be addressed, including which women to include and how to identify and sample these women. Second, a number of integrated strategies were employed to support the letters of invitation. These included targeted media campaigns and the employment of a well-prepared team who were equipped to deal with the volume of enquiries generated by the mailout. In many cases, it was this interpersonal interaction and rapport that was established at the baseline that encouraged women to feel confident and motivated to participate in and continue with the study over the next decade and beyond.

**Acknowledgements**

The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women’s Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

**References**


Dillman DA (1978) *Mail and Telephone Surveys The Total Design Method*, New York, John Wiley & Sons


**EDUCATION RESEARCH BOOKS FROM eCONTENT**

Exclusive distribution of PostPressed Academic titles from September 2007, including:

- *Exceptional Outcomes in ESL/Literacy Education* – by Wayne Sawyer, David Baxter and Paul Brock 74pp ISBN 9781921214127
- *Exceptional Outcomes in Science Education* – by Debra Panizzon, Geoffrey Barnes and John Pegg 84pp ISBN 978-1921214-14-1
- *A Model for Analysing One-To-One Teaching in the Maths Recovery Program* – by Bronwyn Ewing 78pp ISBN 1 876682 77 9
- *Shifting the boundaries: Feminist practices in adult education* – edited by Anita Devos ISBN 1-876682-00-0
- *Teaching Reading and Writing in a Multiliteracies Context: Classroom Practice* – by Annah Healy, 2000 ISBN 1-876682-12-4

EContent Management Pty Ltd, PO Box 1027, Maleny QLD 4552, Australia
Tel.: +61-7-5435-2900; Fax. +61-7-5435-2911; subscriptions@e-contentmanagement.com
www.e-contentmanagement.com
Developing relationships and retaining participants in a longitudinal study

LYN ADAMSON AND CATHERINE CHOJENTA
Research Centre for Gender, Health and Ageing, The University of Newcastle, Callaghan NSW, Australia

ABSTRACT
The strength and success of any research project lies in the participant’s belief that their time and contribution is valued and worthwhile. Developing and maintaining relationships with participants in longitudinal research projects are crucial elements to ensuring the project will meet its aims and objectives. Investing time and resources in the maintenance of the cohort will reward the project with motivated participants, encourage high response rates, lower attrition rates and a representative sample. This paper outlines the methods used in one longitudinal cohort study to develop sustainable relationships with participants.

Keywords: retention; relationships; longitudinal; cohort study

Maintaining contact with participants is essential to the integrity of longitudinal research studies. The success of longitudinal research will reflect the ability of the research team to repeatedly contact their participants and ensure that any non-response from a participant can be explained. This paper describes methods that may be used to assist in the retention of participants in longitudinal research.

In longitudinal research, each and every opportunity for contact with a participant should be valued and conveyed to the participants through a positive and professional attitude. The strength and success of any research project lies in the participant’s belief that their time and contribution is valued and worthwhile. Repeated requests to provide information over long time frames may become irritating and seemingly pointless to participants unless they are provided with tangible evidence that the project is contributing to the wider world of research. Furthermore, the evidence must be meaningful to them. Developing sustainable relationships with participants takes time and money in expenses such as human resources and administrative costs. These costs are not always factored into grant applications, however they may prove to be one of the most worthwhile investments made by longitudinal research teams in achieving their aims and objectives.

Very few populations are static. The reasons for research teams losing contact with their participants are innumerable. The researcher needs to develop a relationship with the participant that encourages them or a relative to advise the researcher of any change of circumstance that might result in a loss of contact between them.

The following methods have been used to develop a successful long-term relationship with
participants in the Australian Longitudinal Study on Women's Health (ALSWH). As the data for this project are collected by postal survey, maintaining current postal and telephone contact is the key to achieving a continual representative response to each wave of the project. The major survey is only sent once every three years. Due to the potential lack of contact for extended periods, the project team has developed a range of strategies to encourage participants to advise changes of circumstances to the research team. Methods have also been developed to help locate participants where contact has been lost. The participants recruited into the ALSWH covered a wide age range from 18 to 75, necessitating development of a variety of retention strategies designed to appeal to each age group.

**Retaining Participants in a Longitudinal Study**

Creating an identity

Presenting a consistent image to participants engenders recognition and loyalty and will in turn support the longevity of the project. The importance of branding to a longitudinal project is difficult to estimate. However, it would seem that branding a research project in a similar way to the branding of any type of merchandise will help embed an image of the project in the minds of the participants and in time engender a feeling of ownership. Creating a highly recognizable image for the project in the form of logos and the consistent style of all communications will enable participants to see at a glance that a specific publication relates to the project. (For further information on the development of the ALSWH logo, see Chojenta, Byles, Loxton & Mooney in this issue).

The ALSWH project team ensures recognition and loyalty in a variety of ways, but principally by monitoring and controlling all communications with participants to ensure that they reflect the overall professional image and standard created for the project at the outset. The project team was mindful of the principal target audience, Australian women, and developed an abbreviated name for the project: Women’s Health Australia. This name, rather than the longer Australian Longitudinal Study on Women’s Health, is used in the majority of communications with participants. As much as is practical the logo and colours are used on every communication with participants.

**Survey packages**

Longitudinal studies are by their very nature repetitive. This repetition may be considered a hindrance to research as participants can grow tired of repeated requests for what appears to be the same information. It is possible to put a more positive spin on the repetitive nature of longitudinal research by regularly communicating results to the participants. Participants look forward to and grow to expect a communication from the project on a regular basis.

Communication with the participants in the ALSWH follows a predictable pattern. Each participant receives a major survey once every three years and a newsletter annually. Over the past decade many participants appear to be conscious of this cycle and call the Freecall number to check if it is a survey year. The survey is mailed out in March and the newsletter in October. This time frame allows the research team time to trace missing participants, prompt participants to return their surveys and assemble the data within the year. The October newsletter is timed to allow the team to track participants without current mailing addresses and remind the age group to be surveyed in the forthcoming year that they will receive a survey in the new year.

All survey packages for both the main, and any additional substudy surveys, (see Helman et al in this issue) are personalised and are printed in an age-appropriate style. The text and font are targeted to the age of the targeted group. For example the surveys for participants in the older age group are printed in a larger font than those for younger age groups. In addition, any communication to the participants in the older age group use the participants’ full title (eg Dear Mrs Jones), whereas in
the younger age group the participants are addressed by their first name within the letter but with a title on the outer envelope. Each survey package consists of a personalised letter of invitation, a survey booklet, a change of details card, and a reply paid envelope for survey return.

In general the formatting of the survey reflects the data collection method however the outer cover of each survey remains largely unchanged ensuring instant recognition by the participant. Each survey clearly identifies the ALSWH project as the sender with the standard logo in the centre of the front page. The outer envelope always carries the project logo in the upper left hand corner to enable the receiver to clearly identify the sender. In addition all major surveys are colour coded to assist the research team to rapidly identify the survey.

Maintaining representativeness

Encouraging and maintaining representativeness in the targeted population is a key element to the success of a longitudinal study. Taking into consideration the needs of participants with culturally and linguistically diverse backgrounds, or disabilities such as vision or hearing impairment is important when planning. Consider the cost of developing a variety of data collection methods that will not compromise the integrity of the results when creating a budget. Although this can be an expensive exercise the benefits of encouraging a diverse sample will be reflected in the quality of the data.

The ALSWH encourages representativeness in the project by preparing separate survey packages and data collection methods for participants with specific needs, while maintaining a sufficiently similar process that data integrity is maintained. Women of culturally diverse backgrounds who have English as a second language receive letters inviting them to make use of trained translators to complete surveys. The translators call the participants shortly after they have received their invitations and establish a time to complete the interview by telephone.

Participants who for a variety of reasons, such as vision impairment or physical inability cannot complete a postal survey are also offered a telephone interview. Once again, these participants are invited to complete the survey by mail and are telephoned by a research assistant to arrange a mutually acceptable time for a telephone interview. Participants who have difficulty completing the survey are encouraged to ask for assistance from family members or carers to help them complete the survey. There is space on the back page of each survey for proxy participants (e.g. family members or carers) to inform the research team that they have assisted the participant. Fields are allocated in both the participant and survey databases to identify people who have made use of proxies (see Adamson & Graves in this issue). The number of women who need to complete the survey by telephone in the older age group is increasing over time as failing eyesight and various physical conditions prevent them completing the survey by hand.

Change of details cards

It is imperative that participants taking part in longitudinal research are encouraged to advise the project of every change of address or circumstance that may result in the project team being unable to contact them. Providing frequent opportunities and encouragement to communicate any changes is the key to maintaining up-to-date contact details. Strategies such as providing Freecall telephone numbers, change of address cards with free return postage or encouraging email correspondence with the project are low cost means of maintaining continual contact. See Table 1 for some examples of the means of staying in touch with participants undertaken by the ALSWH.

The ALSWH has logged over 40,000 changes of address in the first ten years. A change of details card accompanies every mailout. The cards have the ALSWH logo and return address on one side and space for the participant’s details on the other. The wording on the card encourages participants to either return it with their completed
survey or separately by normal post. If participants have no current changes we suggest they keep the card in a safe place and return it should their circumstances change.

In longitudinal research projects the sequence of changes of address may be of more importance to the research team than just knowing the participants last known address for mailing purposes. In the case of the ALSWH project each change is fully dated with day, month and year. This information is recorded in the database and has proved valuable for researchers with a particular interest in the geographical location of participants.

**Reminders**

Mailing surveys to participants is the easy part – getting them back is the challenge. Establishing how often and when participants need to be reminded to send a response back is dependent on the target population. As discussed in Adamson, Young & Byles in this issue, the Dillman (1978) mailing protocol is very comprehensive and suggests several forms of reminders that can be adjusted to reflect the population. Just as in mailing a survey, the presentation and timing of the reminders requires careful consideration. It is easy to overstep the mark and appear to be pressuring participants to return surveys. The importance of piloting each system (see Loxton & Young in this issue) cannot be over estimated. In general, creating friendly, relevant, interesting and eye catching reminders that are timed appropriately will go a long way to encouraging participants to return their surveys.

**Thank you/reminder**

The ALSWH research team use a modified form of the Dillman protocol (Dillman 1978). After mailing any survey a thank you/reminder is mailed out to all participants, thanking those who have returned their survey and encouraging non-respondents, to do so. The reminder is a double-sided flyer that includes the Freecall telephone number and the project email address and website. The reminder mailout is timed to suit the target age group.

**Targeted reminder**

A targeted reminder in a similar format to the thankyou/reminder is sent to non-respondents to the initial survey mailout a further three weeks after the first reminder (see Figure 1). The wording, font and timeline for this reminder is adjusted to suit the three different age cohorts. The targeted reminder details a very short summary of the findings of previous surveys for the particular age group being surveyed and stresses the unique contribution of each participant to the research.

**Table 1: How to stay in touch with participants**

- Provide Freecall or toll free contact numbers to contact the research team. Ensure these are answered during office hours. Provide clear instructions on answering machines with a call back offer.
- Ensure answering machine messages encourage participants to leave their old address as well as the new one. It is surprising, especially in large studies, how many participants will have the same name as other participants and how easily names, numbers and addresses become confused over a phone line.
- Include a change of address card with free return postage with every postal communication
- Include a change of address form on the project webpage
- Respond promptly to all communication from participant

**Figure 1. The reminder process**
**Telephone reminder**

The telephone provides a very effective means of contacting participants to remind them to complete and return their surveys. Undertaking telephone reminder calls can take time and incur large costs. In addition, a recent increase in the activities of marketing companies within Australia has made many people very wary of telephone contact. Trained telephone interviewers who follow well-established protocols are essential to the success of reminder calls. In addition, the need for confidentiality and adherence to relevant privacy laws and the amount of time that calling large numbers of participants can take, should be factored into the data collection time frame and budget. As the relationship with the each participant develops over time, the difficulties encountered in using the telephone as a means to prompt the return of survey, reduces. Researchers should be aware of any standards relating to the contact of persons by telephone. In Australia, the Australian Communications and Media Authority (ACMA) is responsible for establishing these standards. As of January 2007, a draft industry standard for undertaking telemarketing and research activities by telephone is being prepared in response to the Do Not Call Register Act (2006).

The telephone reminder process generally begins twelve weeks after the original survey mailout, but is adjusted according to the timing of the targeted reminder to avoid excessive and unnecessary contact with participants. All telephone operators are research assistants trained to follow the reminder protocols that have been developed to suit each age group and with adherence to ethical guidelines.

The telephone reminder database for the ALSWH project is an electronic record of the outcomes of all phone calls made to each participant during a particular survey mailout. This database is linked to the main participant database. To ensure a participant is not called unnecessarily once completed surveys are returned to the project, the participant’s details are removed from the telephone reminder database. For further information about the design of this database, see Adamson & Graves in this issue.

The outcome of each individual phone call is recorded on the database. This way any interviewer can open the database and see the outcome of each call. As with mailed contact, being mindful of the makeup of the target population ensures positive contact with participants. To maximise opportunities of contact with participants the interviewers adjust the hours they attempt calls. For example, when attempting to contact the participants in the younger and mid age group interviewers work outside of business hours and on weekends. Older women are contacted only in the day time. To ensure confidentiality, strict protocols govern leaving messages on answering machines or voicemail.

Mobile phones and email are a very reliable way of making contact with participants. A simple reminder text message directly to a mobile phone or known email address could be considered when developing strategies for participant contact. The telephone reminder is also a useful way of determining those participants whose contact details are no longer current. Table 2 offers some hints for making successful reminder calls

**Tracking lost participants**

Inevitably some participants in longitudinal research will be ‘lost’. Systems for locating partici-

---

**Table 2: Hints for making successful telephone reminder calls**

- Encourage interviewers to practice the protocols for the study on each other using possible scenarios
- Make sure all interviewers have a good knowledge of the study, including the funding body, the aims, the principal investigators and ethics officer contact details.
- Be aware of any relevant industry standards
- Consider the time zones that you may be calling.
pants should be flexible as communications technology can change markedly over time. Changes in technology that enable new means of possible contact will require amendments to protocols and ethics approvals.

The ALSWH project employs Project Assistants (Warner-Smith, Loxton & Brown in this issue) to assist in locating participants for whom the research team does not have current contact details. All postal returns to senders and non-contacts by telephone or email are recorded in the participant database. Currently all records relating to tracking for each individual participant are recorded on a pre-identified tracking sheet. This sheet is printed from the database and includes the last known address, telephone number and email contacts for the participant and any outcomes from the last telephone reminder. Having access to these details means the Project Assistants do not duplicate calls and can see the results of previous reminders. See Table 3 for some tips and tricks for tracking lost participants.

The following steps are taken by the ALSWH research team in order to reestablish contact with a ‘lost’ participant:

1. **Checking current database details**
   The standard protocol for the ALSWH project team when searching for participants lost to contact is to first check there are no transcription errors in the address. The next step is to use previously untried old telephone numbers and email addresses.

2. **Email and mobile addresses**
   A growing number of participants are supplying email addresses and these are becoming more reliable with time, although email addresses often change and this method is not foolproof. Since 2005 the research team has added separate data fields on the main survey consent form requesting the mobile phone number and email address of each participant. This additional information is proving to be very useful as many participants, particularly in the younger age group, no longer maintain land telephone lines. The past few years has seen mobile phone numbers become more stable in Australia as customers can move from one service provider to another without altering their number.

3. **Stable contacts**
   Subject to privacy laws and institutional ethical requirements, obtaining the name and contact details of a person who knows the participant can be a valuable source of information.

4. **Electoral roll**
   Medical researchers may, subject to ethical requirements, obtain an electoral roll file from the Australian Electoral Commission to aid in their search. The ALSWH currently obtains files that list given names, surnames and current residential address and postal address of electors across three specific age ranges. Information relating to obtaining access to these files and the formats are available from the Australian Electoral Commission (2007).

5. **White Pages telephone directory**
   The online White Pages of Telstra allow internet searches for telephone listings. This search method is becoming less successful as more people choose to not list their number, use mobile phone numbers only, or are listed under their partner’s name. However the initials of account holders often remain the same as they move from location to location, and recording these initials when numbers are initially provided can help locate participants.

6. **Using the address list from the original sample**
   The source of the original list may be worth considering as a further means of tracking lost participants. The original provider may be willing to update your address list by contacting lost participants on your behalf.

**National Death Index (NDI)**
In addition to ensuring data integrity, comparing the database of a longitudinal research project to the National Death Index or register can provide both economic and public relations benefits for the project. Ascertaining that a participant has
died will reduce costs that may have been incurred by devoting time to tracking them. More importantly, not making unnecessary calls to relatives of the deceased participant reduces the likelihood of unintentional distress. The ALSWH is compared with the NDI on an annual basis. For further information relating to the NDI see Adamson and Graves in this issue.

**Newsletter**

Printing a newsletter on a regular basis is an excellent vehicle for maintaining contact and keeping participants informed of the activities of the research team, and the results of the project. Personalise relationships by identifying research team members who are regularly in contact with participants. Distributing the newsletter via the normal means of communication with participants also provides the research team with an opportunity to check the accuracy of participant contact details. If it is returned to sender then the participant must have moved.

The ALSWH project has produced an annual newsletter since its inception. The standard format is an A4 six page colour brochure. The project colours and logo clearly identify the newsletter as relating to the project, however, the content varies. Each age cohort has one page devoted entirely to results and information specific to that age group. The outer pages usually include current events, links to the website, information about personnel and general information that may directly impact on the participants. For example, for the past two years one page has been devoted to information about a new method for linking participant survey data with other national databases.

**Technology**

Websites (see Chojenta, Byles, Loxton & Mooney in this issue) are increasingly important tools for information dissemination, and a useful way of keeping large research teams and collaborators connected to the project. A continually updated website is an equally valuable tool for retaining participants in longitudinal research. Websites create opportunities for participants and stakeholders to keep up-to-date with research output and serve an increasingly important role in the relationship between the participant and the research team.

Many participants in current research projects have the opportunity to complete the survey online. For further information on conducting longitudinal research online see Kraut et al. (2004). It would seem that despite current advances in technology many willing participants do not have the levels of access or levels of expertise to the internet that allow completion of large surveys online.

Currently the ALSWH project is unable to offer survey completion online, however it is certainly something that has been considered. Interestingly enough when the project team surveyed the pilot group of the mid age groups only 33% indicated they would complete a survey online. The main reasons given for not wanting to complete the survey online were lack of access to a computer and lack of confidence in using the internet as well as confidentiality concerns.

The ALSWH website includes specific pages for the participants. This includes provision for participants to complete an automated change of details form that is submitted directly to the project. Research assistants respond to every email to

---

**Table 3: Tracking - tips and tricks**

- Record all contact with participants
- Check for simple errors. Verify the participant is actually lost to contact.
- Collate as much information as possible when in contact with participants
- Note account initials from the white pages directory of the telephone book when recording phone number changes. These generally remain stable.
ensure the participant knows the change to the database has been made. In addition, the participant web page provides an opportunity for participants to keep up-to-date with activities, such as the date of the next survey for their age group, links to published reports, and access to the newsletter and other articles online.

As methods of communication change, new ways of maintaining contact with participants in research projects will develop. Sending mass and targeted reminder messages by SMS text or online messaging to advise participants of a forthcoming survey are just a few of the possibilities. Advances in communication technology brings with it different ethical implications and so future longitudinal projects will need to consider the implications of these innovations.

**Magnets, ID cards, gifts**

In Australia the National Statement on Ethical Conduct in Human Research produced by the Australian Government National Health and Medical Research Council (NHMRC) provides clear guidelines on reimbursing participants taking part in human research. These guidelines are available on the NHMRC website (2007). Some projects offer small tokens to their participants as a way of thanking them for taking part and to promote retention. Possible tokens include sending birthday cards, bookmarks, providing calendars, pens and fridge magnets to remind participants that they are valued contributors to longitudinal research.

The ALSWH project has offered tokens as a way of thanking participants and providing them with the Freecall 1800 number. Two years into the project, fridge magnets in the distinctive green and purple colours with the Women’s Health Australia logo and Freecall number were distributed with the newsletter. These were a very successful means of keeping the project at the forefront of participants’ minds, as several years later participants are still calling the 1800 number with changes of address saying they were prompted by the fridge magnet. In addition, early in the project each participant received a business-sized ID card enclosed within a survey mailout inviting them to write their ID number on to the card and keep it in a safe place.

**Accommodating extenuating circumstances**

In spite of developing comprehensive protocols for the collection of data there are always extenuating circumstances. It is very difficult to account for every personal situation that may arise across the time span of a longitudinal survey. Developing protocols that allow for some flexibility within data collection systems assists with participant retention.

**Temporary residence outside of the data collection area**

Very few research populations are static. As global travel is becoming increasingly commonplace the ALSWH project has found it necessary to develop protocols to accommodate participants who reside overseas temporarily. In the younger age group temporarily residing overseas appears to account for a considerable number of participants who may not have responded to specific waves of the survey. In consideration of this situation, questions relating to the time and dates spent outside of Australia have been included in the latter surveys.

The ALSWH project does send surveys to participants who are living overseas if the participant specifically requests this. The project however cannot pre-pay the return postage as they have arranged for mail drop points at collaborating universities at popular destinations to encourage the return of completed surveys. Collaborating investigators at these universities forward the mail to the project team.

**Not this time**

There are many reasons for a participant not completing a survey in a particular wave. Illness, travel, and lack of time are just a few. It is helpful to develop protocols to accommodate these situations. The ALSWH project records this situation in their telephone reminder database ensuring
that the participant can still remain in the project but will not receive any further reminders.

**Temporary addresses**
Participants may request that a survey be sent to a temporary address. Being able to accommodate these requests will help retain the cohort. The ALSWH project has developed protocols to allow for this. Addresses can be changed temporarily in the database and reminder system created to prompt the Database Manager to restore the original address at a given point in time.

**1800 Freecall hotline**
Encouraging communication with the research team may be aided by providing a means for participants to contact them free of charge. The quickest and most reliable means is to provide a Freecall telephone number. Australian’s telephone system allows free calls from any landline phone within Australia to 1800 numbers. Being able to receive prompt, informed, personalised answers to questions from a member of the research team can help to cement a relationship with a participant. Creating a file that is easily accessed and regularly updated, containing the numbers of useful health information lines and websites, is a valuable tool in dealing with unusual requests or answers to questions outside of the area of expertise of the research team.

The ALSWH project has a Freecall 1800 number. Over the decade the project has been running this has been an outstanding success. This number appears on every communication to participants. The majority of calls to the Freecall hotline relate to the current survey, or a change of address. At the time of a large mailout such as a survey or the newsletter, the 1800 number can receive up to 70 or 80 calls a day. On a day when there is no mailing activity the average is around five calls mostly relating to change of address.

**Missing pages**
Occasionally participants skip pages (turn two pages at once) when completing their surveys. Missing pages in surveys can have serious implications for data quality. At the outset of the project clear guidelines should be created to deal with these circumstances. Reminder prompts at the end of each paper survey to check for missing pages before placing the survey in the envelope may help, however some missing pages are inevitable.

ALSWH has developed protocols for this situation. Research assistants contact participants who have accidentally missed answering at least one full page in their survey in order to complete the data. Although this is costly and time consuming it does provide an opportunity to strengthen a relationship. This situation arises frequently in the older cohort. The implications of collecting extra data this way will also need to be considered by the research team. The ALSWH team concluded that the advantages of collecting missing page data outweighed the potential issues with data integrity.

**Building relationships with participants**

**Indirect contact – media**
Each media opportunity is a potential contact with participants, stakeholders and the general public. Media exposure helps to reinforce the importance of the role of individuals who contribute to research. Developing a good relationship with members of the media will help keep the research project in the public eye. The effect of media on such a large cohort is difficult to assess however the project frequently receives positive feedback from participants such as, ‘I was going to withdraw and then I saw a researcher on the television talking about the project. I forgot how important I was!’

**Unexpected communications**
Longitudinal relationships are reciprocal. Opportunities for communication with participants appear to increase over time as relationships develop and the longevity of the project enhances its integrity. However it is prudent to develop protocols to respond to written requests for help
in areas outside the area of expertise of the project, keeping in mind ethical requirements. This may seem a costly exercise however it will be repaid in loyalty to the project.

The ALWSH project attempts to respond to every written and or verbal enquiry for information. This includes requests from within surveys or any additional correspondence. Many participants, particularly in the older age group, provide additional information by letter. The topics of these communications cover a wide spectrum and can include information about their current health status, health history, childhood, recent events or their thoughts on current affairs. Increasingly, additional information is being emailed to the project, particularly by the mid and younger age group. Research assistants scan the qualitative section at the back of each survey for requests for further information. In all of these situations, a written response is provided, signed by the Project Manager or a Senior Research Assistant. A protocol has been established to ensure the response is appropriate. A draft of the response is created and discussed with the Project Manager or Director. This ensures that potential issues are discussed with at least two members of the project team.

CONCLUSION

This paper has outlined the importance of the development of protocols, procedures, and the role of key staff in maintaining contact with participants in longitudinal research. Investing time and resources in the maintenance of the cohort will reward the project with a pool of motivated participants who have a sense of ownership of the project. This in turn, will be reflected in a high response rate, low attrition rate and a representative sample. The success of the ALSWH in developing sustainable relationships with participants over the past eleven years clearly illustrates the importance of making each participant feel their individual contribution to the project is valued and worthwhile. It is a crucial element to any successful longitudinal research project in achieving its aims and objectives.

Acknowledgement

The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women’s Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

References


Cohort management: Developing and maintaining participant databases in longitudinal studies

Lyn Adamson and Anna Graves
Research Centre for Gender, Health and Ageing, The University of Newcastle, Callaghan, NSW, Australia

ABSTRACT
Creating databases that will support longitudinal cohort studies over extended periods of time is a challenge. The need to record and archive all current and historical activities for each participant can result in large databases. Effectively managing these databases is a key component of achieving successful outcomes for longitudinal studies. This paper describes the methods that have been employed by one longitudinal cohort study to develop functional and flexible databases that will stand the test of time. While not exhaustive these principles provide guidelines that will assist in the creation and maintenance of databases to support a longitudinal cohort study.

Keywords: longitudinal studies; participant retention; data management; cohort studies

INTRODUCTION
With few exceptions data in longitudinal studies only ever increases. Developing a database system to record participants’ details and activities that will last the lifetime of any longitudinal study is a challenge. Creation of a database that is secure, flexible and has the technological power to adapt to every possible change of participant and data collection circumstance that may arise over the length of a longitudinal project requires extensive planning and technical expertise. In this paper the term ‘cohort management’ is used to describe creation and maintenance of all databases relating directly to the identification and activities of participants in a longitudinal study. This paper outlines one set of methods used for cohort management. These methods have been successfully used by the Australian Longitudinal Study on Women’s Health (ALSWH) for the past eleven years.

The resources required to build participant databases – the hardware, software and technical skills – are items that require high levels of funding at the outset of the project. Ideally the central recording system of a longitudinal study will cope with any eventuality without extensive alteration to the underlying architecture. Choosing appropriate well-recognised software that is compatible with established computer systems and that will ensure the necessary levels of security will contribute to achieving this goal, as will hiring staff with the necessary expertise (see Warner-Smith, Loxton & Brown in this issue).

Most longitudinal studies will need to retain information in their participant database that relates to more than just the current contact or per-
sonal details for each participant. It would appear that once the research team has obtained the most recent contact for any participant, the old contact details would become obsolete. However, the experience from the ALSWH suggests that keeping all past contact is advisable. These records may be used for tracking participants who are lost to contact, or for additional research relating to demographics (e.g., population migration). Maintaining and storing historical records in relation to past addresses and contact details can create enormous archives, which in turn have the potential to create both database and paper storage issues for research teams. Well planned and flexible databases and filing systems are crucial to maintaining contact with participants and ensuring data integrity.

**Database structure**

Longitudinal studies are designed to observe the same individuals over a period of time. The need to record and archive each and every activity for every individual being observed can result in large databases. The database needs to provide a sophisticated, stable platform for data storage but still allow data to be input with ease. The ALSWH project uses the application, Microsoft Office (MS) Access as a means of interfacing with the relational MS SQL Server 2000 (MS SQL).

The database tables are created in MS SQL Server and all data is input by using forms created in MS Access (see Figure 1). The ALSWH participant databases store a variety of data relating to the activities of each participant. These include name and address details, and information relating directly to the current and historical participation status of each participant. By creating reports and queries in MS Access the ALSWH can compile the necessary output files for managing the day to day activities of the project. For example, mailing lists for surveys, data files to allow linkage to external data sources and summary participation statistics.

**Database security and confidentiality**

In order to ensure high levels of security for data, backup of all network drives and the SQL server should be undertaken on a daily basis. All the ALSWH databases are password protected. The

---

**Figure 1 Database structure diagram**

- **Input**
  - Change of details
  - Change of participation status
  - Tracking information
  - Logging of survey

- **Storage**
  - Database tables

- **Output**
  - Mailing lists
  - Summary statistics
  - Tracking sheets
  - Lists for data linkage eg NDI
  - Substudy lists
Data Manager creates and supplies these passwords to members of the research team who require access to participant information. In addition, all staff who have access to participant information must sign and keep current a Privacy Protocol agreement (see Chojenta, Mooney & Warner-Smith in this issue). Passwords are renewed periodically to maintain confidentiality.

**Participation Databases**

This section describes the fields that could be included when creating a participation database. It outlines the need to create unique identifiers, and participation status for each individual participant at any time. It describes the levels of accuracy required to maintain address databases to ensure long term follow-up of participants in longitudinal studies and highlights the need to retain historical data relating to participants.

**Identification number – the unique identifier**

Each participant in a longitudinal study requires at least one unique identifier, usually called the identification number or ID. This identifier may take a variety of forms, however it must be in a format that is complex enough to ensure the confidentiality of the participant but simple enough that it can be applied in a practical manner. Although there has been rapid technological change in recent times and data from longitudinal studies is increasingly being stored electronically, there may still be a need to retain paper records for legal or ethical purposes. Creating functional alpha or numerical identifiers that may be practically applied to conventional paper filing systems, in addition to the electronic files, will result in cost and staffing efficiencies.

When constructing identification numbers it is strongly recommended that some basic demographic information about a participant, such as their age and gender, be incorporated into the identification number. This enables the research team to quickly establish basic facts about participants when they call or contact the team. This system could be applied equally to other types of longitudinal research, for example, in clinical trials, by including a number within the identification number that identifies the treatment group to which the participant has been assigned.

Each participant in the ALSWH has been allocated two unique identification numbers:

- the study identification number or ID, known only to data management staff, is a direct link to the identity of participants. ID is the primary key used in MS Access and SQL tables
- the IDalias, which is a second level unique identifier, attached to each individual ID. The IDalias is used in all survey data sent to non-data management staff, investigators and collaborators for analysis.

In terms of cohort management, the ID is the most relevant number and will be the focus of the discussion here. The ALSWH ID is a nine-digit number containing basic demographic information and a check digit. Figure 2 describes the ID, where the first digit identifies the age cohort the participant falls into and is the basis of all filing systems. The first digit A, denotes the cohort: 1 for the younger, 2 for the mid-age, and 3 for the older cohort. The second digit, B, denotes the Australian state of residence of the participant at the time of sampling. The third digit, C, relates to the area of residence within Australia of the participant at the time of sampling, defined as urban, rural or remote.

The middle five-digit number (DEFGH) is a sequential number and simply reflects the number of invitations originally mailed out to each age cohort. For example, 00003 would denote that this participant received the third invitation to take part in the study. This is the unique portion of the ID and is the number that is used within the paper filing systems to maintain documentation in numerical order within each cohort.

The final digit in the ID is the check digit, I. It is invaluable as it prevents the ID number from being incorrectly entered into the database. This number can range from 0-9 and allows the research
team to easily establish that the ID is valid. A worked example for calculating a check digit for a nine-digit identification number is provided in Figure 3. The check digit is calculated as modulus 10 of both the sum of the odd digits when considering the sequence ABCDEFGH (ie A + C + E + G) and twice the sum of the even digits (ie B + D + F + H). The modulus 10 of a number is the remainder when that number is divided by 10.

**Other means of personal identification**

It might also be wise to include other forms of personal identification in the participant database. There are times in a longitudinal study when it may be necessary to confirm a participant’s identity even if the research team has created a unique identifier for each participant. The most common reason for this in surveys collected by mail is a substitution; that is, when a person other than the participant has completed the survey. Requesting other unique forms of identification as part of the data collection procedure can help eliminate substitutions. However, even something as personal as a date of birth can be subject to errors over time, through incorrect data entry or participant error. Alternatives available

---

**Figure 2 Unique identifier**

The 9-digit ID number, ABC-DEFGH-I, as follows:

- A = 1, 2 or 3 age group classification
- B = 1, 2, 3, 4, 5, 6, 7, 8 state classification
- C = 1, 2 or 3 area classification
- DEFGH 5-digit counting code
- I = 0-9 check digit

**Figure 3: Calculating a check digit for a participant identification number**

The formula for calculating the check digit for a 9-digit ID number:

\[ I = 0-9 \text{ check digit, calculated by:} \]

\[ ABC-DEFGH-I \]

\[ I = ( ( A + C + E + G ) + 2*( B + D + F + H ) ) \mod 10. \]

An example of the calculation of the check digit:

If the first eight numbers of the ID is 312-08465, calculate the check digit as follows:

Using \( (( A + C + E + G ) + 2*( B + D + F + H )) \mod 10 \):

\[ (( 3 + 2 + 8 + 6 ) + 2*( 1 + 0 + 4 + 5 )) \mod 10 \]
\[ ((( 19 ) + 2* ( 10 )) \mod 10 \]
\[ ( 19 + 20 ) \mod 10 \]
\[ ( 39 ) \mod 10 \]

divide 39 by 10 and the remainder is 9
so 9 is a valid check digit.

The full ID is then: 312-08465-9
to researchers, depending on the privacy and confidentiality laws of their country, could include unique identifiers such as Medicare or social security numbers.

The ALSWH requests the participant’s date of birth in each survey. This field is checked when the survey is logged and batched. If an error is detected the Data Manager will check across the time points to establish the participant’s date of birth. If there is still doubt other methods will be used to confirm identity, such as demographic information and checking the signature and name on the consent form.

Names
It is not uncommon for people to change names by marriage or for other reasons. If linkage to other datasets, by name and other personal details, is a priority, being aware of variations in participant’s names will assist with obtaining complete datasets. Creating fields in the database tables for first name, second name, third name, surname and maiden name will aid linkage and tracking. Keeping historical records of name changes will allow review if there is any confusion over identity. The ALSWH has a policy of keeping and recording every change and variation of participant names. Although this can create very large files it has proved useful when linking the database to other datasets such as the National Death Index (discussed later in this paper).

Mailing addresses
When conducting longitudinal research by mail, the address is the key component that ensures participants will receive their surveys and remain in contact with the project. If the research team is not solely responsible for mailing out all of the communications with participants, ensuring the address information can be exported from the project database to a format that is readable by mailing house software may avoid unnecessary complications and delays. Many postal systems now give mailing discounts to customers who conform to their regulations and standards. By using databases appropriately, it is possible to automate the system to meet such standards.

Addresses are stored in the ALSWH participant database tables in the following fields: pre-address (eg house names), address (eg number and street), town, state and postcode. Each postal address in Australia is linked to a barcode that serves as a unique identifier. Australia Post offers discounts to customers who have achieved a pre-defined level of accuracy in barcoding their mailing lists. The ALSWH project team uses a software program approved by Australia Post to check postal addresses for their ability to barcode, which helps to eliminate data entry error. In addition, the ALSWH data entry system for entering changes has a built-in check to establish that the spelling of a town name is correct and to establish that the correct postcode and state have been allocated to the address.

Other contact details
The ALSWH enters all possible telephone numbers into fields in the database tables, including: work, home and mobile numbers. These numbers are particularly valuable when tracking lost participants (see Adamson & Chojenta in this issue). The ALSWH research team currently obtains the email address for participants, which is entered into a field in the database tables as a further means of maintaining contact with the participants (see Adamson & Chojenta in this issue).

Participation status
In addition to maintaining the current contact or personal details for each participant, a field in the database tables is used to record each participant’s status in relation to the project. When creating databases to record participant activity for the duration of a longitudinal study, being able to assign a meaningful code for each scenario requires planning. At any one time many scenarios may exist, for example participants may ask to be temporarily withdrawn or to not take part in a wave of the survey. In addition to creating a code for each scenario it may be necessary to create classifications
within a particular code to record extra information relating to the status. For example if a participant asks to withdraw, knowing why they no longer wish to take part may be important for both data analysis and to help prevent further attrition.

The status of each participant in the ALSWH is recorded in the participant status field of the database tables. There are five options: participating, lost to contact, withdrawn, overseas, or deceased. At the beginning of the project everyone’s status was ‘participating’. Now this category includes all those participants who are actively taking part in the project. That is, they have completed the most recent survey, and there is no indication that they have withdrawn, died, become lost to contact or have moved overseas. ‘Lost to contact’ is another straightforward category that includes all participants who can not be currently contacted (ie the ALSWH has no current contact details). More details are required where participants have withdrawn, are overseas or have died, and we discuss these next.

**Withdrawn**

When a participant withdraws from the project their status changes from ‘participating’ to ‘withdrawn’. If an explanation is provided by the participant on withdrawal this information is entered into the withdrawal reason field. This information is valuable for further analysis and reporting requirements. Table 1 shows the currently recorded reasons for withdrawal from the ALSWH project.

**Deceased**

When a participant dies their date of death is recorded and status is changed accordingly. Until the date of death is verified with the National Death Index (NDI) a temporary date of death status is allocated. As ALSWH is a health study, there are a number of other details that need to be recorded in addition to date of death. Obtaining accurate information about the circumstances of a participant’s death is facilitated by use of the National Death Index, which is covered later in this paper.

**Overseas**

Frequently participants reside overseas on a temporary basis and the ALSWH has established protocols to handle this situation. This status is most common in the younger cohort. The ALSWH project team has created four sub-categories to accommodate this situation and to ensure that contact is maintained with the participant (see Table 2).

**Participant Activity and Historical Databases**

In addition to keeping an accurate record of addresses and the current participation status of each participant, maintaining a comprehensive record of each individual’s participation history within the project may be crucial to the data analysis and the integrity of longitudinal projects. Dependent on the nature of the research this could include any number of fields including the date of return of surveys, the dates of clinical

---

*Table 1: Reasons given for withdrawal, recorded in a subfield of the participant status field*

<table>
<thead>
<tr>
<th>Code</th>
<th>Reason for withdrawal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No reason given</td>
</tr>
<tr>
<td>2</td>
<td>No longer interested</td>
</tr>
<tr>
<td>3</td>
<td>Confidentiality/privacy issues</td>
</tr>
<tr>
<td>4</td>
<td>Questionnaire/study issues</td>
</tr>
<tr>
<td>5</td>
<td>Concerns with linkage to other datasets</td>
</tr>
<tr>
<td>6</td>
<td>Caring for someone/has experienced tragedy/trauma</td>
</tr>
<tr>
<td>7</td>
<td>Too ill</td>
</tr>
<tr>
<td>8</td>
<td>Other personal reason</td>
</tr>
</tbody>
</table>

*Table 2: Overseas status, recorded in a subfield of the participant status field*

<table>
<thead>
<tr>
<th>Code</th>
<th>Type of overseas status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Travelling, temporarily not in study</td>
</tr>
<tr>
<td>2</td>
<td>Send survey to an overseas address (partici-</td>
</tr>
<tr>
<td></td>
<td>pant still has Australian contact details)</td>
</tr>
<tr>
<td>3</td>
<td>Send survey to an overseas address (par-</td>
</tr>
<tr>
<td></td>
<td>ticipant has no Australian contact details)</td>
</tr>
<tr>
<td>4</td>
<td>Send survey to Australian database</td>
</tr>
<tr>
<td></td>
<td>address (family members will forward)</td>
</tr>
</tbody>
</table>
events, records of contact by phone reminders, and dates of reminders. In the ALSWH all the tables are relational and both historical and current data can be retrieved for each individual. Ease of input is facilitated by creating forms in MS Access.

**Recording survey activity**

As each wave of the ALSWH project is mailed an input form in MS Access is created to record the return of surveys. The timely entry of this information facilitates creation of reminder mailing lists, telephone reminder lists and additional mailout lists. These records allow output of statistical information related to return rates and reminder effectiveness. It also allows the ALSWH project to track the progress of each survey throughout the data handling process.

**Database applications**

**Using database information in telephone reminders**

There are times in a longitudinal study when the normal means of survey reminders may need supplementing by another method, for example when postal reminders fail to facilitate survey return it may be necessary to use telephone reminders to further encourage participants to return their surveys. Developing MS Access forms to input the outcomes of all reminder calls will minimise costs in both staffing time and telephone charges and ensure that each participant receives the minimum number of calls necessary to prompt the response.

The telephone reminder of the ALSWH project has been created to allow the project team to telephone participants who have not responded to the standard mailed reminders. The principal aim of undertaking a telephone reminder is to actively encourage a participant to complete a survey. By speaking to a participant the project team can establish if the participant has not received the survey, is unwilling to take part in a particular wave of the project, or wishes to withdraw from the project. It is also seen as an opportunity for positive contact with a participant and a means of updating participant details (see Adamson & Chojenta in this issue).

The telephone reminder form allows the operator to view all current demographic information held in the participant database. The form is used to record the status of all calls. The database is updated, and an operator is able to view up to date information as necessary. The form recalls the outcome and action of each call and a notes section allows information that may be useful to the next caller to be recorded. For example ‘participant does shift work, do not call before 4pm’.

**Using database information to track participants**

Developing a system to ensure all possible avenues of contact are explored assists with retention of participants and in finding those who have become ‘lost to contact’ (for more information on tracking ‘lost’ participants, see Adamson & Chojenta in this issue).

Occasionally the project loses touch with participants. Lost to contact advice is normally received by mail being returned unopened to the project or being unable to contact a participant by telephone. These non-contacts are recorded by generating a tracking sheet. The tracking sheet outlines the last known contact details of the participant and the results of attempts to contact them by telephone (drawn from the telephone reminder database). The tracking sheet is given to the project assistants to begin the search for the participant. Once the participant’s current address has been re-established, the participant database is updated with the new information, including the date that the participant was found and the reason why tracking has ceased. Inevitably some participants are unable to be located; in this case the participants are given a participation status of ‘lost’. Lost participant files are periodically reviewed.
Linking participation and external databases using the National Death Index (NDI)

It is an inevitable aspect of a longitudinal survey that a number of participants will die over the course of the project. Ascertaining that participants are deceased is particularly important, as the continual receipt of mail for a deceased person may distress relatives. Lack of knowledge about a participant’s death also incurs unnecessary costs for the project team as they attempt to track participants. In addition, obtaining cause of death data allows the research team to explore the predictors of mortality.

The ALSWH project team receives notification of death in a variety of ways. Approximately half of the death notifications received by ALSWH are provided by relatives, friends, carers or institutions who contact the research team by telephone, letter or email. The remaining half is established by matching participant details to those of the National Death Index (NDI).

The NDI is a database, housed at the Australian Institute of Health and Welfare (AIHW), which contains records of all deaths occurring in Australia since 1980. The data are obtained from the Registrars of Births, Deaths and Marriages in each Australian state and territory. The data includes identifying information and a cause of death code (ICD-10-AM) (NCCH 2004). The NDI is designed to facilitate the conduct of epidemiological studies and its use is strictly confined to medical research (AIHW, National Death Index 2007).

To enhance the quality of the matching process, researchers planning to use the NDI are encouraged to supply to the AIHW as much information as possible including surname, given names, sex, date of birth and date of death (if known), plus a unique identifier for each record, and the last contact date with the participant (AIHW NDI data provision package 2007). The ALSWH matches data with the NDI annually. The application package describing the process for linking participant information with the NDI may be downloaded from the AIHW website.

The files received from the NDI are accompanied by documentation describing the different output files, the methods used for matching and the results. On receipt of the NDI files, the ALSWH conducts a clerical review where each proposed match pair in the output files is either accepted or rejected (AIHW 2007 User-guide to the results). This review is carried out by comparing ALSWH and NDI files in the following categories: dates of birth, all past and current participant names, state of last known residence. In addition, the date of last contact with the participant by the ALSWH is compared with the NDI date of death. Once all checks are completed the participant records are updated with any new deaths or extra information and a request is sent to the NDI for updated lists of cause of death codes from all matched records.

Conclusion
Developing functional, yet flexible databases that will stand the test of time is a challenge. The integrity and longevity of the longitudinal project is however reliant on effective management of the databases that relate to the identification and activities of the participants. This paper has described the methods that have been used by the ALSWH over the last eleven years. While not exhaustive, the principles that drive these methods have been and remain fundamental to the success of the project as whole.

Acknowledgement
The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women’s Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

References


---

### BOOKS & ISSUES ON CULTURE AND HEALTH


This book invites us to rethink the meaning of medicine and life - which are intertwined together. Most significantly, it stimulates our thinking of how to live in a more humane way, and this is the passion that I would like to share with you all. Author

Traditional Chinese Medicine (TCM) is a great treasure of China’s ancient history and culture. Written for health professionals, researchers, social scientists and educators, this book elaborates the embodiment of TCM in the lifespan and complex human dimensions and meanings in Chinese culture.

*Orphaned by the Colour of My Skin: A Stolen Generation Story* by Mary Terszak ISBN 978-1-921348-08-2 viii + 152 pages

I need to emphasise that my time in *The Home of the Good Shepherd* caused me mental trauma, which I feel destroyed my soul. Author

In an invasive, paternalistic, federal public policy environment for Indigenous communities, this book provides an in-depth account of one person’s experiences as a ‘Stolen Generation’ Aboriginal Australian.

Told from the heart, the book speaks in the raw voice of a grandmother reflecting on her life, focusing on her childhood experiences, subsequent perceptions and life stories.

*In Our Own Right: Black Australian Nurses’ Stories* edited by Sally Goold OAM and Kerynne Liddle ISBN 978-0-9757422-2-8 xiii + 120 pages

The intimate, private, and heart wrenching stories told in this book, the first of its kind in Australia, will penetrate the hearts and souls of even the most hardened reader. Told with incredible dignity and humility, each of the individual and deeply personal stories recounted is a powerful testimony to the gross inhumanity and brutal capacity of white people in Australia - colonists who selectively destroy and humiliate, without remorse, the lives and souls of their fellow black Australians. This book provides a powerful catalyst for questioning and calling into question the taken-for-granted humanity of us all.

*Advances in Contemporary Transcultural Nursing* edited by John Daly and Debra Jackson ISBN 978-0-9750436-1-5 xiv + 190 pages

The articles in this collection provide compelling evidence that culture is fundamental to personhood and can be a crucial variable in human experiences around health, illness, recovery and rehabilitative processes. Within the challenges lie opportunities – opportunities to extend and develop practice, to examine the ways that nursing and health care is delivered, to continue to develop our understandings of spirituality and culture, and their relationship to health and well-being.

*Advances in Indigenous Health Care* edited by Eileen Willis, Vicki Smye and Maria Rameka ISBN 978-0-9750436-9-1 xii + 204 pages

The burden of disease and ill-health in colonised Indigenous populations globally is greater when compared to other groups within their nations. Indigenous health status is linked not only to the deprivation and socio-economic positioning, but to the access and use of health services – shredded in politics through Indigenous communities to policy and resource decisions in government. This issue explores differential access to essential determinants of health (such as housing, education, income and nutrition) and health services in North America, Australia and New Zealand and seeks to compare lessons learned.


With a Foreword by Madeleine Leininger, this significant collection addresses the significance of culture-specific care for maintaining the health and well-being of Indigenous people, transcultural nursing leadership & intercultural communication in action, transcultural nursing models (frameworks) for education & practice (combined models, standards, competencies), evidence-based transcultural nursing practice, transcultural nursing ethics, social justice, human rights, and challenges and issues of nursing beyond borders.

---

**Cohort management: Developing and maintaining participant databases in longitudinal studies**

---

**Death Index**


---

**Bibliography**

- Human resources for longitudinal studies: Matching people to skills and tasks
- *Traditional Chinese Medicine: The Human Dimension*
- *Orphaned by the Colour of My Skin: A Stolen Generation Story*
- *In Our Own Right: Black Australian Nurses’ Stories*
- *Advances in Contemporary Transcultural Nursing*
- *Advances in Indigenous Health Care*

---

**Contact**

eContent Management Pty Ltd, PO Box 1027, Maleny QLD 4552, Australia
Tel.: +61-7-5435-2900; Fax. +61-7-5435-2911; subscriptions@e-contentmanagement.com
www.e-contentmanagement.com
Data management involves the planning, management and production of data in a format suitable for researchers to use (see Figure 1 & 2). The products of longitudinal studies are the datasets. Efficient and careful data management will result in datasets that are as accurate and as complete as possible. In addition, effective data management can reduce missing data and minimise data entry error. The final dataset must be in a form that is easy to understand and to use with a variety of statistical packages. Most importantly, data management processes and manipulations must be reproducible and well documented. This paper aims to provide some insight into data management procedures, using the Australian Longitudinal Study on Women's Health (ALSWH) as an example.

Keywords: longitudinal studies; data management; dataset; recoding; derived variables
Data management: The building blocks of clean, accurate and reliable longitudinal datasets

**Figure 1: Data management: Survey design to raw dataset**

- Survey design and planning
  - Question & response design
  - Survey coding
  - Identification numbers
  - Ethics approval
  - Survey piloting

- Paper survey returned to data capture company
- Paper survey returned to project

- Data capture
- Visual review & missing data

- Visual review for missing data
- Data capture

- After data capture or automatic data entry
- Raw dataset

**Figure 2: Data management: Data capture to recoded dataset, storage and archiving**

- Data capture
  - Scanning
  - Manual data entry
  - Automatic data entry

- Storage
  - Digital imaging
  - Paper survey storage
  - Paper survey destruction

- Production of the data set
  - Raw dataset
  - Naming
  - Create formats & labels
  - Raw frequencies
  - Recoding

- Scales and derived variables
  - Recoded dataset
  - Other datasets

- Archiving
  - Data Book
  - Recoded frequencies

- Data distribution

- Documentation: Data Management Manual
- Management: Data Management Group

- Documentation: Data Dictionary Supplement
- Data Dictionary
variables may hold a different type of measure, for example a continuous scale variable, a categorical variable or qualitative response. The information in the variable may be subjective (eg participant’s answers to a question) or objective (eg participant’s results from a clinical test).

In longitudinal studies a dataset is obtained for each wave or phase of the study. These datasets are combined into a single dataset for longitudinal analyses so that there is one record for each case for each wave of the study (Young, Powers & Wheway in this issue). Each variable contains the same kind of information about the case or participant and the information for each variable is comparable over waves eg mental health, alcohol consumption and marital status.

**Survey Administration Method**

There are many different methods of collecting information for datasets and mailed, self-administered, paper surveys are only one of many options. Other methods include surveys administered by an interviewer, email surveys, web page based surveys and clinical results (Loxton & Young in this issue).

**Question and Response Design**

Data quality starts with study design. In the case of surveys (whether interviewer administered or self-administered) there are design issues to be addressed with regard to the wording and layout of questions and responses which will help ensure a smoother process when producing an accurate dataset. These include:

- **Mutual exclusivity**
  For a simple question (one question with a single set of response options from which to choose an answer), the response options should be mutually exclusive so that a participant’s answer cannot fit into more than one category. In multiple choice questions (those with many parts to the response options where all or only some may be selected as responses) mutual exclusivity of response options is not always possible. Some validated scales deliberately ask the same or similar questions in different ways to validate the response information given.

- **Exhaustivity**
  The response options must be exhaustive, in other words the response options available in the question must cover every single possible category. For example in one ALSWH survey a question (Figure 3) was asked about abnormal pap tests. The question was changed in later phases to include an option for ‘I don’t know’ as some participants were unsure of their Pap test results.

- **Double negatives**
  No one would intentionally ask a question with a double negative but when combining the question and response, the result may be a double negative. The following question (see Figure 4), used in a pilot survey, asks about employment. Part A is ‘I do no paid work’. When this is put together with the response option ‘No’ it makes ‘I do no paid work – No’, a double negative. This question was reworded in the main survey. Generally the solution to a double negative is to examine the question and reword it in a positive manner.

---

**Figure 3: Example of a question to illustrate exhaustivity**

<table>
<thead>
<tr>
<th>25 Have you EVER had an abnormal Pap test?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Circle one number only)</td>
</tr>
<tr>
<td>No 1</td>
</tr>
<tr>
<td>Yes, in the last 5 years 2</td>
</tr>
<tr>
<td>Yes, more than 5 years ago 3</td>
</tr>
</tbody>
</table>
Double-barrelled questions
Double-barrelled questions occur when two questions are asked in one. This can lead to confusion where the participant is not sure which part to answer. One possible solution is to split the question into two questions - a lead question (possibly with a skip) followed by a second part to the question. Reworking and rewording of the question might be a better solution, although the wording should not be changed for validated scales.

The ALSWH asked the following double-barrelled question (see Figure 5a and 5b) in a pilot survey. ‘In the past 12 months were there any times you needed to see a medical specialist but did not?’ Participants were trying to be as accurate as possible so they answered that they needed to see a specialist, even if they had seen a specialist. This question was confusing and led to inaccurate responses so it was replaced with a better version for the main survey, as shown in Figure 5b.

**Skip instructions**
Skips are instructions used to direct the participant through some of the questions that follow depending on the answer to the lead question. For example, in Figure 6, if the response to the question ‘How often do you currently smoke cigarettes or any tobacco products?’ is ‘Daily’ then ‘Go to Q47a’. On reading a question that has skips a participant will occasionally follow the skip instruction but forget to answer the lead question. Other problems occur where participants do not follow the skips or become confused by complex skip patterns. On the whole, it is better to avoid skips and especially complex skip patterns.
SURVEY CODING

Once the survey design is complete it is a good idea to classify the survey responses. This involves assigning a description label to each question and a number or code to each response option of the question. Preparing the coding at an early stage can sometimes highlight an error or inconsistency in the wording of a question or the instruction given with it. At this stage the error can easily be corrected. The smoking question (Figure 6) is used as the example again. The first response option ‘Daily’ is coded as a ‘1’, the second response option ‘At least weekly, (but not daily)’ is coded as ‘2’.

IDENTIFICATION NUMBERS

Each ALSWH participant has been allocated two unique identification numbers (code numbers): the study ID and an IDALIAS. These identification numbers are used to control access to identifiers: personally identifying information such as name, image, address, and date of birth. The ID is attached to datasets which contain participant identifiers and the IDALIAS is attached to datasets which contain survey data (see Adamson & Graves in this issue). Records linking the two identification numbers are stored securely as required by the study protocol. As per recently developed protocols for the use of data in health research, (National Health Medical Research Council 2007) datasets from which identifiers have been removed and replaced by a code (eg the IDALIAS) are considered to be re-identifiable.

Each item in the ALSWH datasets is allocated one of two access levels. Items allocated to access...
level A are available only to the ALSWH Data Manager and data management staff, and include the study ID, names, contact details, and date of birth. Level B items are more generally available and are included in datasets with the IDALIAS. Level B datasets do not contain any information that might compromise individual confidentiality. These items are available to study staff and approved collaborators (ALSWH Data Dictionary Supplement 2007a).

**ETHICAL REQUIREMENTS**

Once the survey design is complete the survey needs to be approved by the appropriate human research ethics committee HREC. The National Statement on Ethical Conduct in Human Research is Australia’s primary source of guidance for the conduct of human research (National Health Medical Research Council 2007).

**SURVEY PILOTING**

The importance of pilot testing the survey before use on the main cohort cannot be stressed enough. The ALSWH has separate groups of participants who form the pilot cohorts. These women receive a pilot survey approximately six months before the main cohort is due to be surveyed. Based on the pilot results, questions are omitted, wording is changed (which may require further piloting) and, in general, the survey becomes a better instrument with which to collect the desired data accurately and reliably. For further information see Loxton and Young in this issue.

**VISUAL REVIEW**

Visual review is an appraisal of a paper-based survey after completion by the participant prior to data capture. Visual reviews facilitate efficient data capture and may improve the accuracy of scanning and manual data entry. With the advent of modern scanners and optical mark recognition which alerts the data entry operator to an invalid response, the need to hand edit surveys to avoid scanning errors has been reduced.

Examples of when the scanner might be unable to detect correct responses include:

- the printing bleeds through the page onto a response on the other side
- the participant makes ‘stray’ marks on the page (doodling)
- the participant marks a response and the mark extends into another response
- a participant writes comments in the margin of the survey which stray into the scannable area
- a survey is water-damaged or torn
- the participant marks two boxes for one question and the incorrectly marked response is crossed out
- the participant has not filled in enough of the box for the scanner to detect the marked answer.

Before the advent of current scanning methods, one solution to these issues was to ‘white out’ incorrect responses using correction tape. These days data capture software presents an image of the question causing the problem to the operator and they make a decision as to the correct response.

Visually reviewing surveys before scanning or manual data entry may be necessary for confidentiality reasons. Participants might disclose sensitive or identifying information on the page. Some researchers prefer to have this information ‘whited out’ before data capture and others may prefer to have the sensitive or identifying information omitted from the data output, leaving the original information untouched.

Making these qualitative data confidential requires the editor or data entry operator to be trained as to what data to remove and what to keep. Anything that is screened should have a note added to indicate that this has happened. For example, a participant might refer to her new address in the survey comments. When this is edited or deleted the words [new participant address] might be inserted. Obviously the change of address details must be noted before this happens.
Missing data
Occasionally participants leave one or more questions in a survey unanswered. Missing data of this sort can be detected during the visual review or flagged by the data capture software. The ALSWH project has a policy of phoning participants who have missed one complete page or more and requesting that they complete the missing questions over the telephone. The surveys which have had a telephone follow-up for missing data are batched separately for identification. If the missing data are identified before data capture the questions are completed by the telephone operator directly onto the survey. If the missing data are identified after data capture a database can be constructed to enter the responses to the missed questions into the raw data, or the missing data can be entered into a survey, scanned and used to update the raw data with the extra information.

Data capture and verification
There are several methods of data capture: manual data entry, scanning and automatic data entry. Once the method of data capture has been decided, this may need to be put to tender if it is going to be outsourced. If data capture is not outsourced a database for data entry needs to be created. Suitable database software must be chosen.

Data verification can be defined as the process of checking the accuracy, completeness, consistency and the overall reliability of the data collected. Data verification methods differ depending on the manner in which data are entered.

Manual data entry
For manual data entry a data entry operator reads a paper survey and types the question responses into a computer program. When surveying a small number of participants, manual data capture may be cheaper than scanning. Manually entered data are commonly verified by double data entry using a different data entry operator. Discrepancies are resolved and the data corrected. The industry-accepted error rate for manually entered unverified (single entered) data is 5% when carried out by an experienced data entry operator. It is not advisable to have data entry carried out by an inexperienced operator. It is a common misconception that any junior in the office can enter data. However, like everything else, aptitude and training are required for data entry.

Verified (double entered) data have an accepted error rate of 1%. If resources permit, double entry is recommended. Otherwise partial double data entry is a more economical and industry accepted standard. Data can be double entered for a sample (say 10%) of the surveys or selected sections of each survey (for example questions considered more important than others). Examples of important questions are those which make up physical or mental health scales. If one or more of these question responses is missing the scale score may no longer be able to be calculated.

In 1996, ALSWH phase one data were manually entered and funds did not permit verification. Although the error rate for non-verified hand entered data is 5%, visual review of a sample of the data have shown the ALSWH phase one data to have about a 1 to 3% error rate. Unfortunately, these errors were not random and are concentrated in certain questions, for example, multiple-part questions. This illustrates the importance of designing surveys that are straightforward to complete.

An example of the non-randomness of errors occurred in a survey question with a list of 24 items from a to x (see Figure 7). The participants were asked to circle either a ‘yes’ or a ‘no’ response to each item.

A systematic error in data entry of the long question was revealed. Some participants responded ‘no’ to part ‘i’ and ‘yes’ to ‘l’ but the data were entered as ‘yes’ to i. It is assumed that one of the operators misread the lowercase ‘l’ for an uppercase ‘i’. These errors were corrected in the dataset but verification by a second operator might have picked up the error earlier.
Scanning
Using this method, a scanner is used to detect pen or pencil marks in the response box on a paper survey and to allocate an appropriate, pre-specified code. Since 1998, the ALSWH data have been scanned. The current data company guarantees 0% error for scanned data. Some data (postcode, date of birth and qualitative comments) require manual entry and the numeric fields should be verified.

Once data capture by scanning has begun it is a good idea to verify some of the data. A sample of the data can be checked against the original surveys, photocopied surveys or survey images. The accuracy rate can be compared with the rate that the data company advertises before continuing. Scanned data capture is quicker than manual data capture for projects with large numbers of participants and may be more accurate.

Automatic data entry
Automatic data entry is where the responses are entered directly by the participant into their computer and transmitted via the internet. With the arrival of Internet technology, Internet-based surveys have become popular. The Internet may be used for surveys using email, a web page or by using a combination of both.

The information in this paper deals primarily with managing survey data collected by a pen and paper method. However issues that pertain to dataset production, distribution and documentation will apply to most longitudinal studies and comprise the second half of this paper.

Production of dataset
The production of the dataset involves the following steps:
- naming the dataset
- naming the variables
- creating formats and labels
- combining all raw data files into one file
- checking the raw frequency distributions
- recoding the raw data
- calculating scales and derived variables
- preparing the recoded frequency distributions
- creating qualitative datasets
- creating other datasets.

Datasets are comprised primarily of study items and derived variables that have been calculated using combinations of study items. Datasets are compiled and distributed under strict privacy provisions and in accordance with the data needs of specific research questions. The protection of participant privacy is of paramount importance in the creation of datasets and consequently all identifying information must be removed from the datasets to be used by researchers and collaborators. This identifying information includes the study ID, name, contact details and date of birth. The ALSWH uses SAS programming for creating the datasets for analysis.

Naming conventions – datasets
Research projects, particularly longitudinal projects, have the potential to generate large amounts of data and information about these...
data. Systems of nomenclature, or naming conventions, are used in various fields and situations in order to provide easy access to information. Using a system of naming conventions means that the data user needs to learn only one set of rules to be able to identify various information about multiple datasets and multiple data items within a project. Information that can be conveyed via naming conventions might include:

- **Cohort**: Which group (e.g., cohort or treatment/control) the data relates to.
- **Phase**: For longitudinal studies, which phase of collection the data come from.
- **Survey**: For studies utilising multiple surveys, the relevant survey. Note that this may also be recorded by combinations of other information, such as cohort and phase, as explained below.
- **Survey item number**: Including the item number in the survey to which the data correspond will facilitate the use of the data in conjunction with its original source.
- **Item type**: Whether the item is a survey item, derived variable, validated scale score, or other type of data.

The ALSWH applies naming conventions to both datasets and data items. For datasets, information regarding the cohort, phase, and level of access (available to all researchers or restricted) is provided in the format: WHA_phase_cohort_access level. All datasets names start with WHA which is our abbreviation for Women's Health Australia (the short name of the project) and indicates that the dataset is from the ALSWH project. Phase represents which number survey the data come from, for example, 1 for the first survey of each cohort, 3 for the third survey of each cohort. Cohort is summarised as ‘Yng’, ‘Mid’, or ‘Old’ for the younger, mid-age, or older cohorts respectively. Access level is either restricted (‘A’: containing possibly identifying information) or available to all approved researchers (‘B’: containing de-identified information). Under this convention, the dataset containing data from the fourth survey for older women which is available to all researchers would be named ‘WHA4OldB’ (followed by the file type suffix).

With three age cohorts, almost four survey phases completed to date for each cohort and 300-600 data items in each survey, it is essential to clearly name datasets and variables. At a glance, the age cohort and the survey phase can be identified by the names of both datasets and variables. By looking at the dataset names in Table 1 it is immediately apparent to which age group and survey the data refer.

**Table 1: Numbers of observations and variables in the ALSWH datasets**

<table>
<thead>
<tr>
<th>Cohort</th>
<th>File name</th>
<th>Observations</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>MID-AGE</td>
<td>WHA1MidB</td>
<td>14,099</td>
<td>332</td>
</tr>
<tr>
<td></td>
<td>WHA2MidB</td>
<td>12,338</td>
<td>437</td>
</tr>
<tr>
<td></td>
<td>WHA3MidB</td>
<td>11,200</td>
<td>490</td>
</tr>
<tr>
<td></td>
<td>WHA4MidB</td>
<td>10,905</td>
<td>596</td>
</tr>
<tr>
<td>OLDER</td>
<td>WHA1OldB</td>
<td>12,940</td>
<td>313</td>
</tr>
<tr>
<td></td>
<td>WHA2OldB</td>
<td>10,434</td>
<td>410</td>
</tr>
<tr>
<td></td>
<td>WHA3OldB</td>
<td>8,646</td>
<td>308</td>
</tr>
<tr>
<td></td>
<td>WHA4OldB</td>
<td>7,158</td>
<td>360</td>
</tr>
<tr>
<td>YOUNGER</td>
<td>WHA1YngB</td>
<td>14,779</td>
<td>307</td>
</tr>
<tr>
<td></td>
<td>WHA2YngB</td>
<td>9,688</td>
<td>573</td>
</tr>
<tr>
<td></td>
<td>WHA3YngB</td>
<td>9,081</td>
<td>500</td>
</tr>
</tbody>
</table>
**Naming conventions – variables**

In longitudinal projects it is useful to be able to tell at a glance to which phase and cohort a variable belongs. This is particularly useful when reading or writing programming code for longitudinal studies. For example q29 in the dataset for the first phase of the mid-age cohort needs to be distinguishable from q29 in the dataset for the second phase of that cohort.

For data items, the cohort, phase and survey question number are identified in the name following the format: **cohort_phase_question number**. **Cohort** in this case is represented by y, m or o, for the younger, mid-age and older cohorts, respectively. Phase is again a number representing the number survey that the data come from. **Question number** is represented by ‘q’ followed by the number used in the actual survey that is completed by participants. For example, the data item y2q28a represents the survey item question 28a, in the second survey for younger aged women.

For data items that do not correspond to the survey items, such as derived variables, the cohort and phase are identified as previously described, followed by the variable name abbreviation. For example, m3bmi represents the body mass index (BMI) for participants in the third survey for mid-aged women.

The conventions for naming these variables were made in 1996, when SAS variable names were limited to eight characters. Even though SAS now accepts more characters, other statistical packages were still limited to eight characters until recently. Choose readily recognisable names for variables but these should not be too long, as long variable names become awkward and take up too much space in tables and cross tabulations when doing analyses.

**Formats and labels**

Formats and labels may be called different names when used in different statistical packages. SPSS calls them variable labels (labels) and value labels (formats). Labels assign a meaningful name to a variable. For example, the smoking question shown earlier is q46 in the survey. Its label is ‘Smoking habit at Survey 3 for the Young cohort’.

Formats assign meanings to the codes in the variable. For example, in the smoking question, formats assign words to the codes for y3q46 as follows:

1 = smokes daily  
2 = smokes weekly, but not daily  
3 = smokes, but less than weekly  
4 = does not smoke at all.

Labels and formats are both very useful as a first step if you need some basic information about a particular variable. Keeping formats consistent across phases is helpful to the analyst.

**Combining raw data**

The raw data may consist of several files because the captured data are supplied in stages. These files need to be combined to create the complete raw dataset. A check needs to be carried out on the raw dataset to find out if there are blank or duplicate records for a participant. A blank record will occur if a participant does not wish to complete the survey this time and returns the survey unanswered. A duplicate record may happen if a participant completes two surveys. In this case the first survey completed is usually kept.

**Raw frequency distributions**

Raw frequency distributions are checked for incorrect data, outliers and high proportions of missing data.

**Reencoding skip questions**

Survey questions which have skip (or ‘go to’) instructions will need to be recoded. Validly skipped questions are given a code to indicate the question has been skipped and not missed (eg the smoking question referred to earlier in this paper). The variable for the question affected can be recoded directly or a new variable can be created. If the variable is not recoded there will be an...
inflated number of missing answers and only a proportion of the valid responses will be included, instead of the entire response set.

Skip codes have been applied in the ALSWH data where necessary (ie where some participants have been instructed to skip some items, a dummy code has been entered for these items). The skip value is obtained from the ALSWH Data Dictionary or from the relevant text format file. For example, in the smoking question (q46 in Figure 6), the participants were asked how often they smoke cigarettes or tobacco products. If they did not smoke or if they smoked ‘Less often than weekly’ they had to indicate this and were directed to q48. If the questions that were skipped (q47a and q47b) are not recoded to 0, frequencies will be obtained for the number of cigarettes smoked daily (q47a) and the number of cigarettes smoked weekly (q47b) among a population of women who smoke rather than the survey population.

**Recoding missing data**

As a longitudinal study progresses and questions are asked repeatedly both in different phases and of different cohorts, sets of rules are built to recode missing data in particular types of questions to ensure that the missing data are handled in a standardised way. Missing data have been replaced with plausible values in the ALWSH data wherever possible. It is not recommended to arbitrarily replace missing value with the null value or any other value. The following are some of the rules for recoding that the ALWSH project has devised.

**Algorithm A of recoding – yes/no questions**

Due to the way that some questions are phrased we can sometimes assume that participants have only responded to those items that are applicable to them. This assumption is made for questions with multiple parts and with only ‘yes’ and ‘no’ responses. It is based on the assumption that participants have only responded to the applicable items and have left out the inapplicable answers, even though a ‘no’ response category was available (see Figure 8). This is a fairly conservative approach as all missing items are not simply recoded to ‘no’. The algorithm is explained as follows:

- if all items have missing responses, leave all items as missing;
- if all items have either responses of ‘yes’ or ‘missing’ and there is at least one ‘yes’ response

<table>
<thead>
<tr>
<th>13</th>
<th>Have you consulted the following people for YOUR OWN HEALTH in the LAST TWELVE MONTHS? (Mark one on each line)</th>
<th>yes</th>
<th>no</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Physiotherapist</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>b</td>
<td>Counsellor / Psychologist / Social worker</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>c</td>
<td>A community nurse, practice nurse, or nurse practitioner</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>d</td>
<td>Optician / Optometrist</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>e</td>
<td>Dietitian</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>f</td>
<td>Podiatrist</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>g</td>
<td>Massage therapist</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>h</td>
<td>Naturopath / Herbalist</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>i</td>
<td>Chiropractor</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>j</td>
<td>Osteopath</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>k</td>
<td>Acupuncturist</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>l</td>
<td>Other alternative health practitioner (eg aromatherapist, homeopath, reflexologist, iridologist)</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

**Figure 8: An example of Algorithm A recoding**
then assume applicable answers only were marked and recode missing to ‘no’;
• if one or more items have responses of ‘yes’, one or more items have responses of ‘no’ and other items have missing responses then leave those with missing responses as missing.

Algorithm B of recoding – scaled positive/negative questions
This is an extension of Algorithm A and applies to questions with multiple parts and with response options other than ‘yes’ or ‘no’, for example ‘none’, ‘1-2 times’, ‘3-4 times’, ‘5-6 times’ etc as in q1 from the fourth survey of the younger cohort provided in Figure 9. The algorithm of recoding assumes that only applicable answers have been answered. The algorithm is as follows:
• if all items have missing responses, leave all items as missing;
• if some items are answered positively and no items are answered as the null value, then recode the missing to the null value;
• if some items are answered positively and some items are answered as the null value and other items have missing responses then leave those with missing responses as missing.

Algorithm C of recoding – ‘mark all that apply’ questions, with ‘none of these’ option
This algorithm applies to questions with the instruction ‘Mark all that apply’ and with multiple response options including the response option ‘None of these’ or ‘None of the above’ as in q12 from the fourth survey of the younger cohort shown in Figure 10. These questions are usually given a default value of ‘no’ during scanning. This algorithm is explained as follows:
• if no items are answered and the ‘None of these’ option is given, but is not answered, then assume the whole question was left missing. Recode all items to missing, including the ‘none of these’ option.
• if any items are answered as ‘yes’ and the ‘None of these’ option is answered as ‘yes’, then recode the ‘None of these’ option to ‘no’.

Algorithm D of recoding – logical recoding
This is a logical reasoning approach which uses answers to other questions to recode a missing question. In the example from the fourth survey of the older cohort shown in Figure 11 the participant gave a missing value in (g) ‘Walking more than one kilometre’ and (h) ‘Walking half a kilometre’ then answered ‘Yes, limited a lot’ to (i) ‘Walking 100 metres’. By inference part g and h are recoded to ‘Yes, limited a lot’.

Scales
The ALSWH surveys contain numerous health measurement scales, most of which have been validated in published work. Prior to using these scales in analyses it has been the practice of the study to evaluate the statistical properties of these scales.
12 **In the last 3 years, have you been diagnosed or treated for:**

(Mark *all* that apply)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Gestational diabetes (during pregnancy)</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>Insulin dependent (Type I) diabetes</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>Non-insulin dependent (Type II) diabetes</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>Heart disease</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>Hypertension (high blood pressure) during pregnancy</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>Hypertension (high blood pressure) other than during pregnancy</td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>Low iron (iron deficiency or anaemia)</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>Asthma</td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>Bronchitis</td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>Postnatal depression</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>Depression (not postnatal)</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>Anxiety disorder</td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>Endometriosis</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>Polycystic Ovary Syndrome</td>
<td></td>
</tr>
<tr>
<td>o</td>
<td>Urinary tract infection</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>A Sexually Transmitted Infection (eg chlamydia, genital herpes etc)</td>
<td></td>
</tr>
<tr>
<td>q</td>
<td>Cancer (please specify on page 30)</td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>Other major physical illness (please specify on page 30)</td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>Other major mental illness (please specify on page 30)</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>None of these conditions</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 10: An example of Algorithm C recoding**

14 **The following questions are about activities you might do during a typical day. Does YOUR HEALTH NOW LIMIT YOU in these activities? If so, how much?**

(Mark one on each line)

<table>
<thead>
<tr>
<th></th>
<th>Yes, limited a lot</th>
<th>Yes, limited a little</th>
<th>No, not limited at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>VIGOROUS activities such as running, lifting heavy objects, participating in strenuous sports</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>MODERATE ACTIVITIES, such as moving a table, pushing a vacuum cleaner, bowling or playing golf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>Lifting or carrying groceries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>Climbing SEVERAL flights of stairs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>Climbing ONE flight of stairs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>Bending, kneeling or stooping</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>Walking MORE THAN ONE kilometre</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>Walking HALF a kilometre</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>Walking 100 metres</td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>Bathing or dressing yourself</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 11: An example of Algorithm D recoding**
scales in the survey in which they were first included for each age cohort. The evaluation includes implementation of the scoring recommended in the scale’s validation and an assessment of its utility in the ALSWH data. The main objective is to endorse the use of the scale as recommended. Where this is not supported by the analysis, an alternative is proposed. Users of the ALSWH datasets are strongly advised to read relevant sections of the Data Dictionary and Data Dictionary Supplement (discussed further later in this paper) before using scales in statistical analysis. In particular, the use of a single item from a scale rather than the scale score is rarely appropriate (ALSWH Data Dictionary Supplement 2007b).

**Frequency distributions and data books**

Once the production of the dataset is complete a data book can be produced using the data collected from the surveys. A data book contains frequencies of responses and missing data for each item asked in the survey. These are useful for researchers who are planning analyses and sub-studies and for data management staff and investigators when planning future surveys.

The ALSWH data books contain percentages of responses and missing data for each option of each item presented for four groups categorized by area of residence (urban, large rural, small rural, remote). Overall responses, weighted according to area of residence at phase 1 in order to account for over sampling of rural and remote areas, is also included. Further information can be obtained on the ALSWH website (2007).

**Other datasets**

As well as the survey datasets, some supplementary datasets have been created. Some of these are provided routinely with the datasets (such as the anthropometric variables) and some require a written request for the data (eg demographic variables, Food Frequency Questionnaire data). Information about the dates of death and withdrawal of the participants is available in the participant status file, which is also available as a separate dataset.

The qualitative data are the comments included by participants in response to the final question on each survey (‘Have we forgotten anything?’) and also include responses to other open-ended questions that may have been asked. These data are available for analyses and there is a separate protocol for its use (Chojenta, Mooney & Warner-Smith in this issue).

**Dataset distribution**

The ALSWH data are released by the data manager in accordance with established ALSWH protocols and procedures as described by Chojenta, Mooney and Warner-Smith in this issue.

**Data management manual**

The data management manual describes the step by step process of creating the final datasets for each survey from the raw data. The document can be searched for information on keywords or identification numbers when a query arises. For more information on ‘job manuals’ see Warner-Smith, Loxton and Brown in this issue.

**Data documentation**

The previous sections of this chapter have discussed the practicalities and challenges of collecting study data, capturing this data and producing datasets. Once datasets are finalised, however, another set of issues must be considered. The challenge at this phase of the study is to communicate important information regarding the datasets and variables to everyone accessing the datasets. A formalised system of data documentation or codebook provides uniform information about the study’s datasets, including original survey items, variables and datasets.

As at 2007, ALSWH has collected data from 13 longitudinal surveys, resulting in the creation of 1977 variables in the datasets. Over 300 staff, investigators and collaborators involved with the study access these datasets. Clearly, utilising a system of data documentation is important in man-
aging the task of communicating all information to users.

ALSWH uses several methods to ensure that the study’s datasets are thoroughly documented:
• naming conventions, mentioned earlier, are used for survey items, data variables, and datasets
• a ‘Data Dictionary’ is used to record information about each variable in the datasets
• the ‘Data Dictionary Supplement’ provides information specifying the processes used to derive selected variables.

This section aims to explain each of these methods of data documentation, providing considerations for development of a data documentation system as well as examples from the ALSWH project.

The Data Dictionary
The Data Dictionary is a searchable database containing documentation on each of the variables within a longitudinal study. The ALSWH Data Dictionary uses Microsoft Access. Although there are many programs that will allow the development of a database, it may help to consider the following when choosing the computer program in which the database will be created:
• Can the program do what is required? Will information be a simple list, or will the information need to be retrieved via a variety of methods and presented in various formats? Can the chosen program cope with the amount of information and number of variables in the dataset?
• Is the program robust to corruption? The database may be subject to multiple users, transfers between computers, being attached to emails. Programs that might be considered for creating a database need to be fairly ‘robust’, and be able to tolerate a large workload.
• Is the program accessible to all users? The database may need to be accessed by a wide range of users, who may not have the same computer programs and skills as the database creator. It is best if the database can be created in a commonly used program which is reasonably ‘user-friendly’.
• Can the database be locked to prevent user changes? One way to prevent unintended changes or deletions to databases is to provide users with a locked copy. This ensures that those responsible for maintenance of the database are the only ones able to make changes.

Once the program in which the database will be created has been chosen, it is necessary to plan the database. This may include planning what information will be entered, what format information will be stored and retrieved in, and how the database will be maintained. Some specific issues for consideration might include the following:
• Design: The design of the database will depend on the program in which it is created, the level of information being stored, and the needs of users. Considerations may include how the database will be searched and whether all of the information for a single item will be accessible on the one page, or separated into different categories.
• Quality control: To ensure that information is easily accessible and understood, it is essential to maintain consistency when making entries in the database. Uniformity of entries within the database may be aided by restricting maintenance of the database to only one (or a few) staff, and by documenting the particular conventions and styles that should be adhered to.
• Comprehensiveness: The level of detail and type of information that is stored in the database will vary from project to project, and should be based on the potential uses of the database and needs of the users. Considerations might include sources of items, description and coding of items, inclusion of items in various surveys, and/or whether to provide cross-referencing for similar items.
The Data Dictionary is intended to be used in conjunction with datasets and the original surveys to aid understanding of the data. The following reports can be obtained from the Data Dictionary:

- a list of questions
- a list of formats
- a list of response options
- data specifications for data capture
- a list of additional questions from one phase of a cohort to the next
- a list of deleted questions from one phase of a cohort to the next
- the Data Map (summarises ALSWH data for all surveys and provides a citation list for items that have been included in past surveys. It is a short summary document, to be used in conjunction with the data dictionary).

The ALSWH Data Dictionary database has been designed so that most searches can be completed via a main form. In addition to the main form, the Data Dictionary contains forms listing category codes and names, and forms which collate items for a single cohort (ie young, mid-age or older). Due to the large amount of information stored in the Data Dictionary, these additional forms help to refine the search for topics and data types.

To ensure quality control, naming conventions are used and a particular format of recording information is adhered to within the Data Dictionary. In addition, maintenance of the Data Dictionary is restricted to only a few staff, who make changes according to the documented protocol.

In terms of comprehensiveness, the ALSWH uses the Data Dictionary to aid in understanding data for statistical analyses and to assist survey planning and design, so detailed information about both the scoring for data items and the sources of survey items is required. In addition, similar items are cross-referenced within the Data Dictionary. Cross-referencing assists longitudinal analyses, as items that may have been modified between different time points, and therefore have separate records, can be identified more easily.

Each of the types of information stored in the Data Dictionary is described below.

- Index number: A unique identifier is assigned to each record, of the form category_number. Numbers are allocated sequentially to each record within a category. For example, ALCS-001 is the 1st record in the alcohol (ALCS) category.
- Category: These are broad classes for related variables. Each record is assigned to only one category, although some could theoretically be placed into more than one category. Categories facilitate searches for specific classes of items. For example, a search for alcohol items (category ALCS) would find 15 records from the thousands of records in the database.
- Variable type: This label indicates the type of variable described by the record (see Table 2).
- Access level: Each variable is designated an access level based on the study’s Privacy Protocol (ie Level A for identifying variables and Level B for other variables).

### Table 2: Types of Variable

<table>
<thead>
<tr>
<th>Type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey item</td>
<td>An item from at least one main survey</td>
</tr>
<tr>
<td>Derived variable</td>
<td>A variable developed within the project and created as a function of two or more survey items</td>
</tr>
<tr>
<td>Validated scale/score</td>
<td>A scale/score developed and created as a function of two or more survey items</td>
</tr>
<tr>
<td>Weighting variable</td>
<td>A variable used to adjust for over-sampling of women living in rural or remote areas</td>
</tr>
<tr>
<td>Survey information variable</td>
<td>Variable indicating whether the full or short version of the survey was completed</td>
</tr>
</tbody>
</table>
• Source: A reference for or a description of the source of each variable, for example, a validated scale source reference.
• Description: For survey items, the description is the full text of the item as it appears on the surveys. The full name and a definition are included for derived variables.
• Responses: This field records the possible responses for each variable (e.g. ‘Yes’ or ‘No’). This field will be blank for survey items that require a written response and for some derived variables.
• Response code: This records the coded values for responses (e.g. ‘Yes’ = 1, ‘No’ = 2). This field will be blank for survey items that require a written response and for some derived variables.
• Used in derived variable: This indicates if the variable is used in the calculation of created variables or scales and provides the name, abbreviation and index number for the derived variable.
• Exact item variable name: Indicates survey items or derived variables that match this record exactly, that is, have the same text and responses or scores.
• Similar item variable name: Indicates survey items or derived variables that are similar to the current record but for which either the wording, responses or scoring differ slightly.
• Exact/similar item notes: For similar items indicates where differences between the current record and similar item are found (i.e. within the question, responses, or both). For exact items, this item indicates points to take note of (for example, exceptions to the naming convention for that item).
• Similar item index number: Lists the index numbers for similar items, to assist in cross-referencing within the database.
• View supplement: When this indicator is marked there is further information about the variable in the Data Dictionary Supplement. It also provides a hyperlink to the Data Dictionary Supplement on the ALSWH website.
• Info for data users: This field provides a hyperlink to information and tools for data users on the ALSWH website.
• View surveys and data: Data users are encouraged to use the Data Dictionary in conjunction with the surveys. This internet link provides easy access to the surveys for data users. In addition, the link provides access to the ALSWH data books, which provide frequency statistics for each of the main surveys.

Other data documentation
Most research projects collecting quantitative data will make use of validated scales and other derived variables within their datasets. However, scales and derived variables need to have detailed information recorded about them in order to convey the full meaning of the scale and its derivation to users. Such information might include the source of the variable, the derivation process, and/or the statistical programming code for the variable. Often, this will be too much information to be recorded in a database. One option is to provide a supplementary document that provides detailed information about such variables.

The ALSWH Data Dictionary Supplement provides information about derived variables and scales within the ALSWH datasets. These variables are validated and endorsed by members of the ALSWH Data Management Group (described below). The Data Dictionary Supplement aims to provide a transparent account of the derivation of the variable, as well as recommendations for use. The following information is included in the Data Dictionary Supplement for each derived variable or validated scale:
• Variable information: Description and naming convention for the variable.
• Person/s associated with preparation: Staff members who have been involved in the development of the variable.
• Date variable is endorsed by Data Management Group: This ensures that the variable has been passed through the Data Management Group.
• Recommendations for use: There will have been a lot of work completed on a derived variable prior to its inclusion in the datasets.
Recommendations ensure that data users have access to the outcomes of this work.

- Development process: This section provides a detailed summary of the principles and process behind the development of the variable. It also includes relevant statistical program code (in this case, SAS code) for the derivation of the variable.
- References: Any references that were used to inform the development of the variable are listed.

**Data Management Group**

As data and statistical issues and problems associated with a longitudinal study have arisen, the ALSWH has found it helpful to form a group to discuss and resolve these issues. This group consists of the project manager, data managers, data assistants and statisticians.

The issues that this group may address may include:

- how to handle errors in the data
- what to do about missing data
- how to recode new questions
- how to handle transition variables
- how to manage validation of scales
- how to monitor the progress of the survey and the production of the main and additional datasets.

**Data Storage**

When running a longitudinal study, do not underestimate the amount of storage which will be required. There are two types of storage:

- physical storage - paper surveys, participant consents and other written documentation
- electronic storage
  - data files, data dictionaries etc
  - digital imaging files.

**Physical storage**

Some ethics committees require that paper surveys are stored for a minimum number of years. Other ethical monitoring bodies may be satisfied that the images of surveys are stored instead of the paper copies.

Initially the ALSWH project had a requirement that the paper surveys from participants were kept for the duration of the study (20 years) plus another five years. In 1996, between 40,000 and 45,000 surveys from the three age groups were received. The number of paper surveys to date has reached 120,000. These paper surveys are stored in several large compactuses, and can be located by the box, batch and sequence number.

**Electronic storage – data**

Apart from the issue of paper storage there is also the storage of electronic data. The survey data have to be stored in several forms for each survey:

- one or more raw data files as they have been received after data capture
- a single raw data file after the individual files have been combined
- a data file after any recoding has been carried out
- different forms of datasets depending on the statistical program that is used for analysis
- subsets of the data that are not required by all analysts (eg qualitative data).

All datasets need to be labelled in a way that makes it easy to distinguish what the data consists of. When talking about electronic data storage it should go without saying that all datasets should have a backup system in place in the event of hardware failure or accidental deletion of the data.

**Electronic storage – digital imaging**

Technological developments mean that surveys and consent forms can now be archived electronically as digital images. These take up a lot of hard drive space and it is convenient to store these on external hard drives for quick access or on CD or DVD. The data in the digital images is confidential and must be stored securely.

Depending on the data capture software used, digital images can be stored as .tif or .jpg files. The image files should be given their identification number (ID) as their name and if they are jpeg images a page number is tagged onto the ID.
All images for a survey must be stored in a folder marked appropriately for easy identification.

In order to ease the retrieval of data images, a database application may be created. It is often necessary to go back to the actual survey or images to check outliers and other anomalies in the data.

Precautions need to be taken before destroying paper copies of surveys. These precautions include making a random check of the survey data against their images. This checks the accuracy of the dataset to the data in the images. It is also advisable to check the names of the image files against the ID numbers in the dataset. This confirms that there are images, for all survey pages, corresponding to the ID numbers in the dataset.

ARCHIVING
The datasets are obviously one of the most important assets of a longitudinal study. Ensuring the security of the data is assisted by off-site archiving. A further advantage of archiving is that it makes the data available for future users subject to certain conditions. People and projects come and go and there is a need for some permanent storage which is well-documented.

The ALSWH project team has a policy of archiving data with the Australian Social Sciences Data Archive (ASSDA) at the Australian National University on an annual basis. Each year we archive the most recently completed dataset, and may re-archive earlier datasets if there have been substantial changes. Information on the procedure for archiving data with the ASSDA is available on their website.

CONCLUSION
This paper has described the process of building a longitudinal dataset, using the Australian Longitudinal Study on Women’s Health as a case study. The aim of data management is to produce clean, accurate datasets with minimal missing data that are easy to understand and analyse, and are reproducible. These aims are achieved by using the corner stones of good survey design, accurate data capture, careful recoding and detailed documentation.

Acknowledgements
The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women’s Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

References


de Vaus D (2002) Surveys in social research, 5th edn, Allen & Unwin, Australia


Working with longitudinal data: Attrition and retention, data quality, measures of change and other analytical issues

ANNE YOUNG, JENNIFER POWERS AND VIRGINIA WHEWY
Research Centre for Gender, Health and Ageing, University of Newcastle, Callaghan, NSW, Australia

ABSTRACT
Longitudinal studies are important because they can help provide answers to questions about cause and effect, although their complexity leads to a number of challenges for the researcher. By their very nature longitudinal studies may continue over a long period of time and/or have many data points and therefore good documentation of procedures is essential. In addition, it is important to develop dynamic databases that reflect the current status of participants in the project and to develop protocols for dealing with inconsistent or missing responses over time. This paper provides some guidance about these issues as well as information about longitudinal data structure and ways to summarise and display the information obtained from longitudinal studies.

Keywords: longitudinal studies; panel studies; research methods

This paper will outline some of the issues that arise in data analysis that are specific to longitudinal studies. This paper will not focus on complex statistical techniques for analysis, as there are many good reference textbooks available. Rather, this paper will prepare the researcher for the types of challenges faced when preparing longitudinal data for analysis, some quality assurance activities that can be established and some tips for how to present effectively some of the process and outcomes data that are collected.

The paper is divided into sections that cover the role of personnel who may be responsible for data preparation and analysis and how best to facilitate communication and dissemination of key decisions and reports; the importance of developing dynamic databases that reflect the current status of the project and its participants, particularly regarding the monitoring and the impact of loss of study participants; dealing with inconsistent responses over time and missing data; and how to quantify and summarise the transitions that occur over time in variables of interest. The paper concludes with a discussion of the importance of considering the sample size required to test hypotheses of interest, given that the data from longitudinal studies are correlated within subjects, and finally some tips about longitudinal data structure and how to summarise very complex results over time into reader friendly and scientifically sound publications and presentations.
DATA MANAGEMENT AND STATISTICAL PERSONNEL
While some longitudinal studies employ large numbers of data managers and statisticians, smaller longitudinal studies may rely on only one data manager and statistician. Regardless of the size of the longitudinal study and the number of personnel employed, it is essential that information relating to the integrity of the data is carefully documented. By their very nature, longitudinal studies continue over a potentially long period of time and/or cover multiple time points. During the study, personnel may change (Warner-Smith, Loxton & Brown in this issue) or the reasons that certain actions were taken may not be recorded. Clear documentation allows a relatively seamless transition between new and retiring personnel as well as being an important part of the written corporate memory.

Where the longitudinal study is large, and particularly if it is geographically diverse, it is useful to establish a formal group that is responsible for documentation and dissemination of information relating to data integrity and guidelines for analysis and presentation of the study data. In the Australian Longitudinal Study on Women’s Health (ALSWH), membership consists of statisticians who analyse the data, data management staff who are involved in the preparation of the datasets (Graves, Ball & Fraser in this issue) and investigators who oversee the strategic planning of analysis activities and quality assurance. Meetings of the group, either in person or by teleconference, provide a regular forum for discussion of all aspects of data management, including:
• assessing the validity, reliability and responsiveness of new survey items
• agreeing on a set of procedures for evaluation of scales
• documenting the validity and reliability of new scales included on surveys
• endorsing all decisions made by the group
• disseminating this information to the wider research group.

The statistics group also provides statistical advice and collaborates with investigators and associates in substudies and subsidiary analyses, as requested.

RECRUITMENT AND RETENTION INFORMATION
Longitudinal or panel studies may have different designs, including cohort studies, rotating samples and supplemented longitudinal studies. In a cohort study, an attempt is made to measure characteristics of the same people at all time points throughout the study. Rotating longitudinal studies aim to reduce respondent burden by surveying participants at some but not all surveys (on a rotating basis). Supplemented longitudinal studies, such as some household surveys, supplement their sample after baseline by adding participants, for example household members who move in after the initial sample was recruited. One advantage of continuing to recruit in this way is that the sample tends to be more representative of the current profile of the population (in this case, households in the community). The sections that follow in this paper relate particularly to cohort studies.

Regardless of the type of longitudinal study, it is important to record information about the success of the initial recruitment process and the extent of ongoing participation in the study. These details are essential for writing papers and reports on the study. Recruitment details should include a description or definition of the population from which the sample was drawn, how the sample was drawn (for example individuals or households randomly selected, over-sampling by age group, area of residence), selection criteria and what proportion of people who were selected agreed to participate in the longitudinal study. These figures can be used to estimate a variety of response and retention percentages (Tooth et al. 2005).

Where possible, the reasons people selected in the sample do not become part of the study should be collected. For example, a person may
be selected who was not eligible for the study, i.e., did not meet the selection criteria or was incapable of completing a survey. Some individuals may not receive the invitation to participate, for example, where the survey was returned to sender due to an incorrect or outdated address or where a telephone number was disconnected. Response rates can be calculated after adjustments for these cases. For example in the ALSWH, the following people were considered ineligible:

- males; not within the age group; not living in Australia; deceased
- too ill to complete e.g., Alzheimer’s, stroke
- unable to complete due to blindness, literacy etc
- in hospital, having surgery
- language difficulties and declined an interpreter
- not permitted to participate (as deemed by doctor, husband or other family).

To examine the representativeness of the cohort, participants (people who agreed to be in the study) should be compared with non-participants (people who actively refused or did not respond) on demographic and other characteristics where possible. The characteristics available for such comparisons will be highly dependent on the information available for the populations from which the samples were drawn. For example, if a random selection is made from people on the electoral roll in a specific area, the age and gender of participants and non-participants could be compared. However, if telephone directory listings were used as the sampling frame, such comparisons would not be possible. As neither of these sampling frames covers the entire Australian population, the ALSWH used the Australian national health insurance database (Medicare) that includes almost all citizens and permanent residents.

By collecting and analysing information on eligible participants and non-participants, the researcher can report on whether the initial sample was representative of the population of interest and whether the study findings might be affected by non-response. For example, comparison of ALSWH participants with women of the same age in the 1996 Australian Census showed ALSWH women tended to be more educated and more likely to be born in Australia (Brown et al. 1998). While demographic comparisons were made, comparisons of health behaviours such as smoking were not possible.

For subsequent waves of the study, you need to determine whether an individual was able to participate (i.e., was still alive and had not withdrawn from the study), and if so, whether the participant was a respondent or non-respondent in the current wave of data collection. Reporting of response rates for longitudinal studies is inconsistent in the literature. Should people who have died or withdrawn from the study (i.e., actively declined any further participation) be included as non-participants at all subsequent surveys? This would seem unreasonable as they are no longer sent surveys and hence are unable to participate in subsequent surveys. Regardless, it is important to record the dates of death or withdrawal from the study and whether an individual could be contacted during the current wave. Then a decision can be made about whether to present respondents as a percentage of all initial participants or whether to report retention rates, i.e., respondents as a percentage of current participants, with a decision whether to include or exclude those who could not be contacted during that wave. There is debate about how to deal with participants who could not be contacted, as some of these may be passive refusals but in cases where their contact details are incorrect, they may not have received the invitation to participate. In such cases it could be argued that these participants should be excluded from the denominator as they are unable to participate in the survey. Ideally several response rates can be presented by calculating percentages with and without those who could not be contacted.

Some of the different types of response rates are shown in Table 1 for women in ALSWH.
When there are high percentages of women who cannot be contacted, such as in the ALSWH Younger cohort, retention is higher when no-contacts are excluded. A lesser effect is seen among the Mid-age cohort, while for the Older cohort, deaths and withdrawals have the biggest impact.

When reasons for non-response are recorded it may be possible to use this information to reduce the non-response in later waves of the study. For example, extensive tracking may find more of the ‘lost’ individuals. More effort spent on involving the participants in the study may result in fewer withdrawals or non-response. Information about non-response also needs to be collected so that the potential effects of attrition or loss to follow-up can be examined. Problems caused by attrition often relate more to the nature of attrition than to the amount. Random attrition only affects the efficiency of the estimates, but non-random attrition can result in unacceptable biases, particularly if it is associated with unobserved individual characteristics (Little 2002; Schafer & Graham 2002) or if the research question relates to a group that is highly susceptible to attrition (Deeg 2002).

Within the framework of the ALSWH, the following information is collected for women at each survey and is used to assess the impact of attrition:

- date the completed survey was returned
- eligible to complete the survey (yes/no)
- attrition status, categorised as
  - respondent
  - person contacted but did not return survey
  - person not able to be contacted
  - mental or physical incapacity to complete survey
  - withdrawn from the study
  - deceased
- date withdrawn and reason for withdrawal
- date deceased and whether from the Australian National Death Index or not.

Details about a participant’s death may be obtained from relatives, friends and/or linkage with the Australian National Death Index (Adamson & Graves in this issue). The Australian National Death Index was assessed as a tool for establishing the vital status of older women in the ALSWH and found to be effective (Powers et al. 2000). Middle names were found to be very important in correctly identifying women who had died.

### Longitudinal Data Issues
Measurement errors are a concern in longitudinal studies where the main aim of such studies is the measurement and analysis of change. Longitudinal data provide the opportunity to identify responses that are not consistent or plausible over time because responses from the same individuals can be linked and analysed. For example, a

<table>
<thead>
<tr>
<th>Table 1 Percentage response and retention at Survey 3 for women in the Australian Longitudinal Study on Women’s Health</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age at Survey 1</strong></td>
</tr>
<tr>
<td>a. Number of participants at Survey 1</td>
</tr>
<tr>
<td>b. Number of respondents at Survey 3</td>
</tr>
<tr>
<td><strong>Response percentage = b/a*100</strong></td>
</tr>
<tr>
<td>c. Number of current participants at Survey 3</td>
</tr>
<tr>
<td><strong>Retention = b/c*100</strong></td>
</tr>
<tr>
<td>d. Number of current participants at Survey 3, excluding those not contacted</td>
</tr>
<tr>
<td><strong>Retention with no-contacts excluded = b/d*100</strong></td>
</tr>
</tbody>
</table>
respondent may report smoking at one survey but at a later survey report to the contrary by saying they have never smoked. It would not be possible to identify such errors in cross-sectional data.

There may also be a ‘learnt component’ within longitudinal surveys whereby participants become more accustomed to the questions over time and are likely to find them easier to answer. In addition, there is evidence in the literature that individuals become ‘conditioned’ by repeated surveys i.e. responses given in one survey may be influenced by those given in previous surveys or questions may be answered differently when participants know the issues being studied (Anstey & Hofer 2004). As these factors might result in errors in the measurement of change, inconsistencies in responses over time need to be investigated so that the errors can be minimised. One strategy to minimise errors is to look for inconsistencies when cleaning the merged longitudinal data and where possible, edit the data according to agreed protocols. In the ALSWH the majority of women provide data that seem valid at a particular time point but the data collected across surveys needs to be checked for consistency.

For example, in the ALSWH women are asked to report their height and weight at each survey (either in imperial or metric measures). While most of the survey data are scanned, the heights and weights are entered by hand. Hence there is the potential for both participant error and data entry error. This raises several questions about how to deal with potential measurement error, such as ‘are extreme values for self-reported weight or weight change correct?’ The values of reported weight, height and body mass index (BMI calculated as weight in kilograms/height in metres squared) across Surveys 1 to 4 were examined in the ALSWH. Extreme values for BMI were observed infrequently but fairly consistently across surveys suggesting the reported values were true though extreme. On checking a sample of the survey forms, data entry errors were found to be rare. Furthermore, data published from longitudinal studies of free-living populations of women where body weight was objectively measured showed the true range of weight changes which women experience. A wide range of values of real changes were reported and so the option of deleting the values for weight (and BMI) when an extreme change is reported was rejected.

What should be done about self-reported changes in height? While a reduction in height may be expected among the Older women as they age this is not expected among Younger and Mid-age women. An examination of reported heights at the first four surveys showed some inconsistencies. A decision was taken by the ALSWH data management group that a single estimate for height should be made for each of these women, based on all the available data. Rules for the estimation of height were based on the consistency and magnitude of difference in the heights reported at the four surveys. Examples of self-reported heights for Younger women, the rule applied and the single estimate of height are shown in Table 2.

### Table 2: Examples of estimated height based on self-reported heights for Younger women at four surveys

<table>
<thead>
<tr>
<th>Survey</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Estimated height</th>
<th>Rules for estimation of height</th>
</tr>
</thead>
<tbody>
<tr>
<td>165.2</td>
<td>165.2</td>
<td>165.2</td>
<td>165.2</td>
<td>165.2</td>
<td>165.2</td>
<td>If height is the same at two, three or four surveys, estimated height is set to the measurement that is consistent</td>
</tr>
<tr>
<td>163.5</td>
<td>167.0</td>
<td>163.5</td>
<td>–</td>
<td>–</td>
<td>163.5</td>
<td>If height is only available at one survey, leave as is</td>
</tr>
<tr>
<td>164.6</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>164.6</td>
<td>If height is only available at one survey, leave as is</td>
</tr>
<tr>
<td>166.0</td>
<td>161.8</td>
<td>162.1</td>
<td>163.4</td>
<td>–</td>
<td>163.3</td>
<td>If the difference between any pair of heights is less than 5cm, then take the mean</td>
</tr>
<tr>
<td>162.1</td>
<td>163.6</td>
<td>168.7</td>
<td>–</td>
<td>–</td>
<td>162.8</td>
<td>If the difference between any pair of heights is less than 5cm, then take the mean</td>
</tr>
<tr>
<td>175.2</td>
<td>–</td>
<td>178.4</td>
<td>–</td>
<td>–</td>
<td>176.8</td>
<td>If the difference between any pair of heights is less than 5cm, then take the mean</td>
</tr>
<tr>
<td>183.4</td>
<td>176.5</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>missing</td>
<td>If the differences between all pairs of heights are greater than 5cm, then height is not estimated</td>
</tr>
</tbody>
</table>
The examples above dealt with time-varying measures (weight) and time invariant measures (height in younger and mid-age women). One limitation of such systematic checking of data (or other methods such as replacing missing values with the mean or median of non-missing values, or using a ‘last value carried forward’ method) is that any subsequent statistical analyses using these values will systematically underestimate the variability of the data.

For some variables being measured it is reasonable that the responses will change over time and many patterns of responses are possible, making it difficult to apply checks on validity of the responses. For example, younger women’s aspirations about being married when they are older are unlikely to remain exactly the same from the ages of 18 to 30 (Figure 1). In the cross-sectional results for the three surveys, a consistent figure of 86% to 88% of younger women aspired to be married at age 35 (measured when they were aged 18 to 23 years at Survey 1, aged 22 to 27 years at Survey 2 and 25 to 30 years at Survey 3). However, only 75% of younger women at all three surveys aspired to be married at age 35 and conversely, only 4% said at all three surveys that they did not want to be married.

Some time varying measures are likely to be directional. For example, the highest educational qualification awarded can only increase whereas other measures are more variable in the way they can change. A person can start or quit smoking at any time although starting smoking is less common and quitting smoking is more likely with increasing age.

Differential data checking and editing may lead to biased results where some groups of errors are easier to detect than others. For example if a person has been diagnosed as having diabetes, it is a chronic condition and can reasonably be assumed to be an enduring condition, even when a response may be missing or inconsistent on this item.

**Analytical issues**

An advantage of longitudinal studies over cross-sectional studies is that they can help provide answers to questions about the dynamics and determinants of individual behaviour and provide a means of testing hypotheses about cause and effect. Longitudinal studies are essential to measure gross changes in the behaviour of populations such as the changes in alcohol consumption over time. In addition, such studies can distinguish between permanent and transitory characteristics of a risk factor such as drinking alcohol at dangerous levels. Is this behaviour consistent over surveys or is it evident only at a single survey? How does this behaviour relate to other events?

---

**Figure 1** Younger women’s aspirations about being married at 35 years of age, recorded at Survey 1 in 1996 (S1), Survey 2 in 2000 (S2) and Survey 3 in 2003 (S3)
While longitudinal data are very useful, their very nature makes data coding and analyses more complex. Firstly how do you deal with refinement or changes in questions over time? Will the response options be compatible over time? For example, when asking about consultations with medical practitioners over a particular time period, the response options (number of visits) are sometimes presented in ordinal categories to minimise respondent burden. However as people age they are more likely to respond in the higher categories and so the upper category may need to be expanded into more groups to distinguish the high users in subsequent surveys. In this case it is possible to be backwardly compatible but in other situations where the wording of the question is changed for greater clarity, it may not be possible to compare responses over time.

Transitions or changes in categorical variables over time may be difficult to monitor and describe. Even where a categorical variable only has two levels, it does not take many waves before there are a large number of response patterns. For example if an individual is simply classified as a smoker (S) or non-smoker (N) at each of four surveys, there are potentially 16 transitional categories, such as SSSS, SSSN, SSNN, and so on. While it might be appropriate to look at all categories in a younger age group where there is considerable change in smoking status, in an older age group it is likely to result in small numbers of participants or no participants at all in some categories. Depending on the sample and the research question, you might want to restrict the number of categories to those that you think are important for your research question. For instance, if you were interested in the effects of smoking (measured at four time points) on an outcome, where the length of time since quitting smoking was important, then categories could be defined as: never smoked; ex-smoker at Survey 1 (that is, had been a smoker but quit by Survey 1 and did not resume during the period of observation); ex-smoker at Survey 3; ex-smoker at Survey 4; or a smoker at Survey 4, hence reducing the analysis of a large number of potential patterns of responses to six categories.

Another important factor to consider is missing data as it may affect the size of the sample available for analysis and the generalisability of the results. Data may be missing for individual items within a survey or for an entire wave(s) due to death or withdrawal or when a person cannot be contacted or does not respond to a particular survey. Missing data is problematic if it is likely to bias the results of the analysis. For example, withdrawal from the longitudinal study will result in missing data at subsequent waves. If withdrawal is related to an unmeasured characteristic such as smoking and also to the outcome of interest then the results will be biased. However if the missing data are predictable from other non-missing data, then possible bias may be correctable.

Analysis techniques for datasets with missing data include:

- using multiple imputation to ‘fill in’ missing data and analyse the multiply imputed ‘complete’ datasets;
- inverse probability weighting or weighted estimating equations provides extra weight for each person with complete data in proportion to the number of similar people with missing data.

These techniques and models assume that the missing data are predictable from other non-missing data. If this assumption is not plausible, more complicated models such as selection and pattern-mixture models may need to be used (Singer & Willett 2003). These and other methods are being developed to deal with deaths and withdrawals due to frailty (Diehr et al. 2001; Bowe et al. 2006).

**Preparing longitudinal data for analysis**

As with any analysis, several factors need to be considered when preparing to analyse longitudinal data such as the research question, the sample
size required to detect a meaningful difference and the data structure. When analysing data from longitudinal studies, we may want to test two different types of hypotheses:

- are the mean scores for each group equal?
- are the rates of change in the response variable over time for each group equal?

If the outcome variable is continuous, then the research question of interest may be comparing means from two or more groups at each point in time. Assuming the variable is continuous, the rate of change over time can be estimated as the slope of a straight line drawn through the two or more time points at which the response is measured. So, differences in rate of change can be measured as differences in slopes. Figure 2 demonstrates the difference between the two situations using a graphical representation of each hypothesis, using fictitious data.

Perhaps the most frequently asked question concerning the study sample is ‘What size sample do I need?’ The answer to this question is not always simple, and is influenced by a number of factors such as the research question, the measure of change, the variability of the data and the anticipated attrition over time. In longitudinal studies, the same people are being measured at each time point. Hence the responses of each individual are expected to be correlated over time and the standard sample size formulae must be revised to incorporate this correlation between repeated measures. Sample size can be calculated by applying a series of mathematical formulae (Diggle et al. 2002; Hedeker & Gibbons 1999; Twisk 2003)

Before the longitudinal data can be analysed, the data from each survey or wave needs to be merged for each individual. Usually the merged dataset will contain a single line of data (a record) for each individual. For longitudinal data, the data need to be arranged so that there is one record for each individual for each wave of data collected. The following example shows how to structure data for a continuous outcome (mental health) and two categorical explanatory variables (alcohol consumption category and marital status) measured at four time points. Mental health (MH) is measured here on a scale from 0-100, where higher scores represent better mental health. Alcohol consumption (A) has four levels: non-drinker, occasional drinker, moderate drinker and heavy drinker. Marital status (M) is coded as married or not married.
The original dataset takes the conventional format with a single line per individual, with one column of data for each variable at each survey (Table 3). Hence there are three lines of data for three individuals (one record per individual). The longitudinal dataset consists of 11 lines for the three individuals, with one column for each variable and a variable indicating the corresponding wave of the survey (Table 4). ID 2 did not respond to wave 2 and so there is no line of data for wave 2 for ID 2.

### Table 3 Mental health (MH), alcohol consumption (A) and marital status (M): extract of dataset in conventional format

<table>
<thead>
<tr>
<th>ID</th>
<th>MH1</th>
<th>A1</th>
<th>M1</th>
<th>MH2</th>
<th>A2</th>
<th>M2</th>
<th>MH3</th>
<th>A3</th>
<th>M3</th>
<th>MH4</th>
<th>A4</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>84</td>
<td>2</td>
<td>1</td>
<td>68</td>
<td>3</td>
<td>1</td>
<td>60</td>
<td>3</td>
<td>1</td>
<td>60</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>76</td>
<td>2</td>
<td>1</td>
<td>96</td>
<td>2</td>
<td>1</td>
<td>80</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>80</td>
<td>3</td>
<td>1</td>
<td>88</td>
<td>4</td>
<td>1</td>
<td>88</td>
<td>4</td>
<td>0</td>
<td>88</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

When summarising and presenting results from longitudinal studies, it is important to consider who the audience for the report or presentation will be and their expectations. For example, the conventions used to prepare a publication for a peer-reviewed journal will be very different from the style of writing used for a newsletter to the study participants (Chojenta, Byles, Loxton & Mooney in this issue). The media release that may accompany a new publication will require a concise summary of the main findings but with a reader-friendly style. The challenge when summarising the analysis and results from a longitudinal study is to convey the breadth and depth of the data, as well as the careful analysis, while keeping the ‘take home’ message clear. In particular, who is the audience for this publication or presentation and what are their expectations and interests?

There are many ways to present information through graphs and diagrams - the challenge is to know the most appropriate types of graphs and charts for your data. The data may be summarised as means and standard deviations of continuous variables (such as heights or weights) measured for several groups across several time points. Alternatively you may have categorical measures where you are more interested in the percentages of subjects with an outcome of interest over time (such as survival rates).

The following example uses a mosaic plot to show changes in physical activity between two surveys. Figure 3 shows that only 36.4% of the Younger women were categorised as doing ‘moderate to high activity’ at both surveys, with 16.3% remaining in the ‘low activity’ and 2.6% remaining in the ‘no activity’ category at both surveys. Almost a quarter (22.3%) of the Younger women moved into a higher physical activity category during this period (from ‘no activity’ to either ‘low activity’ or ‘moderate to high activity’, or from ‘low activity’ to ‘moderate to high activity’) while 22.4% moved into a lower category of physical activity. These descriptions of change in activity levels for individuals would not be possible in multiple cross-sectional studies.

**Presenting Longitudinal Data**

When summarising and presenting results from longitudinal studies, it is important to consider who the audience for the report or presentation will be and their expectations. For example, the conventions used to prepare a publication for a peer-reviewed journal will be very different from the style of writing used for a newsletter to the study participants (Chojenta, Byles, Loxton & Mooney in this issue). The media release that may accompany a new publication will require a concise summary of the main findings but with a reader-friendly style. The challenge when summarising the analysis and results from a longitudinal study is to convey the breadth and depth of the data, as well as the careful analysis, while keeping the ‘take home’ message clear. In particular, who is the audience for this publication or presentation and what are their expectations and interests?

There are many ways to present information through graphs and diagrams - the challenge is to know the most appropriate types of graphs and charts for your data. The data may be summarised as means and standard deviations of continuous variables (such as heights or weights) measured for several groups across several time points. Alternatively you may have categorical measures where you are more interested in the percentages of subjects with an outcome of interest over time (such as survival rates).

The following example uses a mosaic plot to show changes in physical activity between two surveys. Figure 3 shows that only 36.4% of the Younger women were categorised as doing ‘moderate to high activity’ at both surveys, with 16.3% remaining in the ‘low activity’ and 2.6% remaining in the ‘no activity’ category at both surveys. Almost a quarter (22.3%) of the Younger women moved into a higher physical activity category during this period (from ‘no activity’ to either ‘low activity’ or ‘moderate to high activity’, or from ‘low activity’ to ‘moderate to high activity’) while 22.4% moved into a lower category of physical activity. These descriptions of change in activity levels for individuals would not be possible in multiple cross-sectional studies.
When measures are taken at more than two time points a different graph or plot may be necessary. Figure 4 shows differences in the mean out-of-pocket costs for a general practice (GP) consultation according to age and/or area of residence. We have data for seven years, for three age groups of people living in urban and non-urban areas. The out-of-pocket costs have been categorised into five groups. Figure 4 illustrates the changes over time in the costs for mid-aged women living in an urban area.

The following example shows trends over time in the mean general health of Mid-aged women according to their alcohol consumption, measured at four time points. These women consistently reported being non-drinkers, occasional drinkers, moderate drinkers or heavy drinkers at four time points from 1996 to 2004. General health is a subscale of the SF-36 health-related quality of life scale and is measured on a scale of 0–100, where higher scores represent better general health. Figure 5 shows mean general health over time for the four groups of women, along with 95% confidence intervals.

**WHAT INFORMATION SHOULD BE INCLUDED WHEN REPORTING THE RESULTS OF LONGITUDINAL STUDIES?**

Although longitudinal research began many years ago, structured reporting requirements – a guide as to what details should be included in publications so that the reader can assess the quality of the study – are still being developed. A useful checklist of 33 criteria related to threats to the internal and external validity of longitudinal research has been developed (Tooth et al. 2005). Topics covered include the study rationale, population and generalisability of the findings, recruitment, data collection, biases and data analysis. For example, did the study specify:
Working with longitudinal data: Attrition and retention, data quality, measures of change and other analytical issues

![Bar chart](image)

**Figure 4** Displaying trends over time in a categorical variable

- the number of participants at each wave?
- reasons for loss to follow-up?
- missingness of data items at each wave?

When describing the analysis methods, good practice would be to:
- state by name the method of analysis

- use longitudinal methods of analysis – those analysing change in outcome over two or more time points and which take into account that observations are likely to be correlated
- take account of loss to follow-up, confounders and missing data.

![Line chart](image)

**Figure 5** Mean general health (SF-36 subscale, range 0-100) of mid-aged women, by alcohol consumption, 1996 to 2004

- - - - non-drinker
- - - - occasional drinker
- - - - moderate drinker
- - - - heavy drinker
There are a range of data analysis issues that need to be considered when analysing longitudinal data and there are many textbooks that cover these subjects (Diggle et al. 2002; Singer & Willett 2003; Twisk 2003). For example, when a longitudinal study has many participants, the study may have enough power to detect small differences over time or between groups that may not be of practical importance. What is the magnitude of the change in the variables of interest that would be considered of practical or clinical importance?

**Conclusion**

A key requirement in managing longitudinal studies is the careful consideration of statistical issues such as those outlined in this paper, discussion of options for dealing with inconsistencies, missing data and so on, and the documentation and dissemination of decisions to the wider group of collaborators who are using the data. As the longitudinal study unfolds there will be additional complexities to resolve and the importance of good data management and reporting mechanisms cannot be overstated.

**Acknowledgements**

The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women’s Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

**References**


Conducting substudies in a longitudinal research project

JENNY HELMAN, DEBORAH LOXTON, LYN ADAMSON, ANNA GRAVES AND JENNIFER POWERS
Research Centre for Gender, Health and Ageing, The University of Newcastle, Callaghan NSW, Australia

ABSTRACT
Longitudinal studies often include substudies which involve collection of specific and more detailed data from subsets of study participants. The longitudinal study framework adds methodological strength to these substudies through enabling sampling of individuals with exposures or outcomes of interest, and through retrospective and prospective access to longitudinal data. However, while there are many advantages to these studies, there are also a number of potential disadvantages. Here we describe some of the considerations when designing and approving a substudy and some of the procedures to be applied to ensure that the substudy runs well and has minimal impact on study participants.

Keywords: longitudinal studies; substudies; piggyback studies; supplementary studies; research methods

Participants in longitudinal studies sometimes take part in additional surveys or research activities in the periods between each major survey. In the context of this chapter, the term substudies (also known as side studies, nested studies, piggyback or supplementary surveys) refers to studies that are conducted with samples drawn from a longitudinal study. Substudies may be cross-sectional or longitudinal. Substudies normally arise when a specific issue that requires further investigation is identified in the longitudinal study, and may involve either a simple random sample or a sample selected from specific subgroups. A classic example of a substudy is the nested-control study, where participants with an outcome of interest and matched controls are selected from the cohort. Nested cohort studies likewise select participants with a particular exposure of interest and a comparison group who do not report the exposure. The substudy may collect more detailed information regarding the exposure (e.g., dose, duration, type) as well as more information about the outcome (such as validation of self-reported diagnoses). Substudy design and sampling issues are not the focus of this paper although a sound understanding of these issues is essential to ensure valid results are obtained. (For example, see Beaglehole, Bonita & Kjellstrom 2007; Gordis 2004; Rothman & Greenland 1998.)

This paper will outline the advantages and disadvantages of inviting participants in longitudinal studies to undertake additional research activities and describe the protocols and methods used by the Australian Longitudinal Study on Women’s Health (ALSWH) when undertaking substudies.

It is clear that there are advantages and disadvantages in inviting participants to take part in...
substudies (see Table 1). Additional research will add value to the longitudinal study by allowing the in-depth investigation of specific topics and the extension of research methods (eg combining qualitative and quantitative research methods, or collecting observational data). For instance, the ALSWH surveys have used quantitative methods to measure depression and depressive symptomology among the Younger cohort. Restrictions on the length of the ALSWH surveys (see Loxton & Young in this issue) meant that it was not possible to determine how young women cope with depression. A substudy used quantitative and qualitative questions to investigate this topic in more detail.

Very few studies appear to have investigated the effects that substudies have on subsequent response rates in longitudinal research. The findings of those studies have been contradictory. For example, results from one study indicated a decrease in attrition associated with substudy participation (Gades et al. 2006), while another study found no association between attrition and substudy participation (Deeg et al. 2002). These findings offer some support for the idea that reasonable requests for further information will not adversely affect the overall response rate. However, the lack of available information in this area points to the need for more research before conclusions can be drawn about the impact of substudies on retention and attrition in longitudinal research.

Indeed, it can be argued that more frequent contact may increase participants’ sense of engagement with the research and may help to minimise non-response in the future. Since substudies are an additional contact opportunity, it is possible that substudies contribute to retention, as was found by Gades et al. (2006). In addition, reporting the substudy results to all participants in the longitudinal study assists in emphasising the value of the study overall, and the value of each additional contribution. Results from substudies can also generate media interest, which in turn offers an opportunity to raise the public profile of the longitudinal study. Media coverage has been known to encourage retention of participants (Adamson & Chojenta in this issue), to contribute to research transfer and may generate further interest in the longitudinal study.

Substudies also offer an excellent opportunity for research and higher degree students. Such substudies benefit students, by offering access to high quality data and the opportunity to partici-

<table>
<thead>
<tr>
<th>Table 1: Advantages and disadvantages of conducting substudies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td><strong>For the longitudinal study</strong></td>
</tr>
<tr>
<td>Value adding to the longitudinal study</td>
</tr>
<tr>
<td>Increased participant contact with longitudinal study team</td>
</tr>
<tr>
<td>Raise public profile of longitudinal study</td>
</tr>
<tr>
<td>Interest in substudy topic may generate further interest in longitudinal study</td>
</tr>
<tr>
<td>Provides research opportunities for students</td>
</tr>
<tr>
<td><strong>For the substudy collaborator</strong></td>
</tr>
<tr>
<td>Collaborator has access to longitudinal data</td>
</tr>
<tr>
<td>Ability to select on specific characteristics, exposures or outcomes</td>
</tr>
<tr>
<td>Administrative support</td>
</tr>
</tbody>
</table>
Conducting substudies in a longitudinal research project

Substudies in a longitudinal research project are distinct advantages to any researcher who conducts a substudy. For example, access to longitudinal data that predate the substudy, the ability to select participants with particular characteristics, exposures or outcomes, the opportunity to obtain longer-term follow-up of health outcomes through the longitudinal study, and the support of an experienced longitudinal research team.

The advantages of substudies must be carefully balanced against the possible disadvantages. Arguably the most important of these is the potential for increased participant burden. To mitigate participant burden, protocols and guidelines for undertaking substudies can be developed. The ALSWH has determined that the most important principle in the conduct of substudies is that the longitudinal study cohorts are not considered as convenience samples. Therefore, substudy researchers must justify the need to draw a sample from participants in the longitudinal study.

Ensuring participants are not overburdened by being asked to complete too many surveys within a given timeframe may encourage retention in the longitudinal study. For example, the ALSWH does not conduct substudies of the Younger cohort in the same year that the Younger cohort is surveyed for the longitudinal study. Providing an opt-out clause with each invitation to take part in a substudy gives participants a chance to refuse an invitation without jeopardising their existing relationship with the longitudinal study.

Substudies may impact on the longitudinal study administrative team in a number of ways, such as increased telephone calls from participants, mail returned to sender and the tracking of non-respondents to the substudy. It might be expected that returned to sender mail and tracking would be simply shifted to a different time, however the participants may move between longitudinal and substudy surveys. For example, when women in the Younger cohort were aged 22 to 27 years, 20% had moved house twice and 28% had moved house three or more times in the previous three years and these were the women who could be contacted! Hence women who had moved may need to be tracked during the substudy as well as the longitudinal survey. Although these activities impact on staffing costs, they could be seen as means of assisting retention by increasing positive contact between participants and the longitudinal study staff (see Adamson & Chojenta in this issue).

Another possible disadvantage for substudy researchers is the potential for selection bias. To evaluate the suitability of the longitudinal sample as a sampling frame, the researcher should consider the representativeness of the longitudinal sample with respect to the population he/she wishes to study, attrition since the commencement of the study (see Young, Powers & Wheway in this issue) and the impact on substudy sampling that will be imposed by longitudinal study protocols. Although not easily measured, the relevance of the substudy topic to the participant may influence their decision to participate and therefore also affect sampling. For example, in an ALSWH substudy of women in their twenties, women were asked about the strategies they used to cope with depression. The overall response rate was 76%, but women who had recently become depressed or who had overcome depression were more likely to respond (80% and 78%) than those who had always been depressed (72%) or who had never been depressed (73%).

In order for a longitudinal study team to assess the relative merit of a proposed substudy, a formal approval process will be of some assistance. The efficiency of substudies will be improved for both the collaborator and the longitudinal study team if there are guidelines for substudy procedures. Guidelines can include protocols for the preparation, costing and sample selection process; administration and documentation of the substudy; data collection procedure; and winding-up process. In the following sections we outline the guidelines that have been developed by the ALSWH for the conduct of substudies.
1. APPLICATION FOR A SUBSTUDY

A substudy application process allows a thorough evaluation of the feasibility of the substudy to be made by the longitudinal study team and the applicant. In addition, a formal application provides transparency in longitudinal study decision-making processes and keeps all members of the team informed about the research activities of others.

Applications to undertake substudies will be subject to the procedures and protocols of the longitudinal study. The person (or people) who are applying to undertake a substudy, subsequently referred to as the collaborator(s), will benefit from informally investigating whether their proposal is feasible within the framework of the longitudinal study. Questions that should be considered by the collaborator include:

• Will the proposed research add to knowledge in the field?
• Is it necessary to conduct a substudy of a longitudinal study to answer the research question?
• What research has already been done by the research team in this area?
• Is a similar research project already being investigated?
• Is the longitudinal study an appropriate sampling frame for the substudy? (ie consider selection bias, number and characteristics of available participants)
• Will the time frame fit with the existing longitudinal study procedures? For example, will timing clash with existing procedures?
• Are the proposed research methods suitable for the longitudinal study sample? For example, an online survey may not be suitable for a substudy using a sample of elderly people
• Is the available funding sufficient to conduct a substudy?

It is beneficial to have a nominated liaison investigator from the longitudinal study who can assist with answering these questions and with planning and facilitating the substudy procedure. Negotiation between the collaborator and the investigator can help ensure that the substudy will run on time and within budget. The investigator can also explain substudy guidelines and other longitudinal study resources that might be of assistance to the collaborator.

An application for a substudy is similar to a grant application and in fact, may be a preliminary step to writing a grant application. A substudy application should include a rationale, objectives, methodology (including participant selection criteria), and analyses, as well as a timeline and budget for the substudy. Open communication between the collaborator, the investigator and the longitudinal study team is likely to produce a thorough application, which then forms the basis of the substudy procedure.

It is useful to establish an administration system to track substudy applications. This keeps the longitudinal study team and the collaborator informed about the progress of substudy proposals. As substudy proposals are received, they can be allocated a project identification number and then sent on to the longitudinal study team for evaluation. This system ensures that decisions about acceptance or rejection of the substudy are made in a timely manner (within a few weeks). If modifications or clarification are required, a revised proposal may be requested before a final decision is made about the application.

As with any large scale research project, the potential exists for collaborators to submit applications that overlap with other researcher’s interests or plans. However, in the experience of the ALSWH, it is unlikely that two proposed substudies would be identical. Where proposals are similar, collaborators have been encouraged to work with each other in order to answer their own research questions as part of a single substudy. Agreements between collaborators are brokered by an appropriate ALSWH liaison person appointed by the ALSWH Steering Committee.

Substudy approvals are contingent on the substudy researchers having appropriate funding. It is sometimes the case that a substudy proposal is sub-
mitted prior to funding being granted. For example, researchers might submit a substudy proposal at the same time as they apply for a research grant. Where a grant is unsuccessful, the approval will be withdrawn. The ALSWH procedures clearly state that substudies must be funded, and without such funding the substudy cannot proceed.

If the substudy is approved, the collaborator will need to formalise their relationship with the research team as required by the longitudinal study. Examples of these requirements might include a Memorandum of Understanding (MOU) and a Privacy Protocol. An MOU states the terms and conditions for the substudy procedure and is a formal, signed agreement made between the collaborator and the longitudinal study. A Privacy Protocol usually includes outlines of National Privacy Principles and reference to the privacy protocols that are in place at the longitudinal study. The Privacy Protocol agreement is also a signed formal document. (For more information, see Chojenta, Mooney & Warner-Smith in this issue.)

Within the ALSWH, the following procedures are provided to assist collaborators with the substudy application process.

The collaborator(s):
• discuss their proposal with an ALSWH investigator to establish feasibility
• hold a meeting with ALSWH staff to determine timelines, budgets and the selection criteria for the substudy sample
• use the Expression of Interest (EOI) form for substudies to complete their proposal and apply for approval to conduct the substudy (See Chojenta, Mooney & Warner-Smith in this issue)
• include preliminary copies of substudy materials (eg. surveys, interview schedules)
• send the completed proposal to the ALSWH Project Administrative Officer
• sign an MOU and Privacy Protocol upon approval of application
• request relevant existing longitudinal study datasets.

The ALSWH longitudinal study team:
• allocates a project identification number
• establishes an administrative trail for the project
• assists with the development of the substudy procedures
• distributes the proposal to the longitudinal study team for evaluation
• evaluates the application
• conveys the decision about the proposed substudy to the applicant
• provides relevant existing ALSWH datasets upon receipt of the signed Memorandum of Understanding and Privacy Protocol.

2. PREPARATION OF THE SUBSTUDY

Upon approval of the substudy application by the longitudinal study team, preparation to carry out the substudy begins. Generally, collaborators must seek approval from relevant Human Research Ethics Committees (HREC) before contact with participants can proceed. In this section we describe the procedures used by the ALSWH in preparation for a substudy.

Development of the materials

The collaborator and the ALSWH team develop the substudy materials. Materials include those that will be sent to participants as well as an information sheet about the study for members of the longitudinal study team who will have contact with the participants. In addition, verbal protocols and/or a training program for staff who might have contact with participants are developed. For example, answers to questions that are likely to be asked by participants are prepared, and staff members are briefed as to the appropriate responses.

Another consideration when approaching participants is the potential to cause alarm among groups of interest that are defined on the basis of research data rather than clinical diagnosis. The invitation to participate in ‘sensitive’ substudies may give participants only limited information about the reason for their selection.
For example, in the substudy about coping with depression as mentioned above, the following introduction was used. ‘Most of us have times when we feel happy and other times when we are sad. We are interested to find out what strategies you may have tried to make yourself feel better when you have felt sad, and whether you think they worked for you.’

All study materials that will be sent to participants are produced in the ALSWH style with which participants are familiar, a tactic that assists in the retention of participants (Adamson & Chojenta in this issue) and is likely to encourage participation in the substudy. A substudy information sheet pro forma and an example of a substudy consent form are provided in Figures 1 and 2 respectively. Once finalised and approved by relevant HRECs, the longitudinal study team arranges for printing of materials that will be sent to participants.

Dear *(Title) Name,*

Thank you for your continued participation in the Women's Health Australia project. Your involvement is highly valued by all of us. Once again we would appreciate your help. We would like to invite you to join an important, additional study.

**Who is conducting the study?**
Researchers working on the Women's Health Australia project at…………. .

**What is the study about?**
This year, we plan to ……………………

**Why was I chosen?**
You were randomly selected from the Women’s Health Australia database.

**What do I have to do?**
If you are willing to help, please...

**Do I have to participate?**
Your participation is completely voluntary. If you do not wish to take part you are under no obligation to do so. You are entirely free to discontinue your participation at any time without reason or to decline to answer particular questions.

**What if I do not want to participate in this study?**
If you would prefer not to participate, or if you have any questions, please call Women’s Health Australia on our FREECALL number 1800 068 081. You can also...

**Are there any risks?**
While we do not anticipate any problems, if you do become upset by any of the questions you can call Women’s Health Australia on our FREECALL number 1800 068 081 or ...

**How confidential is the information I give?**
All the information you provide to us will remain confidential and will only be made available to the researchers. Your name and address will not be released and the results of the research will only be published in a form whereby you cannot be identified. All information will be stored at the University of Newcastle in locked filing cabinets or on a password protected computer database. To save you re-answering questions that you have answered in previous surveys, we will link your responses from this survey to your answers from previous surveys.

**How can I find out about the results of the study?**
The results of the study will be published in our annual newsletter. *(and/or other)*

**Who can I contact about the study?**
If you have any questions about the study, please call Women’s Health Australia on our FREECALL number 1800 068 081. Alternatively you can contact one of the researchers.

**NAME AND CONTACT DETAILS**

Thank you for your consideration.

---

1 Women’s Health Australia is the alternative name of the ALSWH, and is used in all participant communication.
Consent

Validation of survey measures in an older cohort

- I have carefully read and understood the Participant Information sheet and voluntarily consent to participate in this research study.
- I also consent to the researchers ‘matching’ the information provided in this survey with that given in the previous surveys so that any changes in my health can be noted.
- I consent to being interviewed on the telephone in more depth about my answers to these questions. (Please tick)

YES NO

_____________________________    ________________________________
Signature                                        Date

What is your date of birth? 19
(Please write date in boxes)

Help us keep in touch!

We plan to survey women in your age group again in two years’ time. Sometimes we lose touch with participants. It would be helpful if you could give us details of a relative or friend who will be able to help us find you.

Name
__________________________________________________________________________

Address

_____________________________   P’code  ___________________________

Phone (Home) ______________________   Relationship to you ____________________________

Please sign above and send the completed survey back to us in the envelope supplied as soon as possible.

We will detach the consent form and store it in a separate locked room.

Figure 2: Example of a substudy consent form (this example includes a request for an interview by telephone)
The timeline
A timeline gives an estimate of how long the sub-study will take to complete and helps to keep the project ‘on track’. It also allows the collaborator to have a reasonable idea of when the information sought will become available.

An estimate of the time taken for the individual tasks in each step of the ‘order of procedure’ is made. Critical events that might cause delays are identified. Tasks or events that are dependent on a previous task or event and the time lag between them are identified. The placing of these tasks and events in a chronological order creates a timeline. Before finalising, the timeline is compared with the ALSWH project timeline, to avoid conflicts of events and staff resources. Clashes with ALSWH events are resolved by altering the substudy timeline or by assigning more resources to that task. A delay in one task can be counteracted by faster action in another. Of course, the quicker each step is completed, the earlier the data become available.

ALSWH uses Microsoft Project 2003 © to develop a timeline that begins at the time the substudy Expression of Interest (application) is developed. Briefly, this software produces a chart based on individual tasks and an estimate of the time taken for each task. Called a Gantt Chart, it displays the tasks, their overlap and the time taken for each group of tasks. It has the advantage of easy adjustment if revision of time estimations is necessary. Table 2 shows the time frames that are allowed by the ALSWH for common substudy tasks.

A timeline can not prevent every ‘hiccup’. In a recent, large ALSWH substudy, long delays in its planning caused a clash with the mailing of the annual ALSWH Newsletter to participants. It was decided to postpone the mailing of the substudy targeted reminder so as not to overload those participants with ALSWH mail. On the upside, the ALSWH Newsletter acted as a reminder and as a result, fewer targeted reminders needed to be sent.

### Table 2: Time frame employed by the ALSWH for substudies involving surveys

<table>
<thead>
<tr>
<th>Time</th>
<th>After</th>
<th>To complete the following action:</th>
<th>By whom</th>
</tr>
</thead>
<tbody>
<tr>
<td>14 days</td>
<td>Receipt of EOI</td>
<td>Consideration and decision making about the proposal</td>
<td>ALSWH</td>
</tr>
<tr>
<td>10 days</td>
<td>Approval of EOI</td>
<td>Complete and sign MOU and Privacy Protocol</td>
<td>Collaborator</td>
</tr>
<tr>
<td>4 weeks</td>
<td>Approval of EOI</td>
<td>Apply for ethics approval, supply survey materials/interview materials, letters to participants etc</td>
<td>Collaborator</td>
</tr>
<tr>
<td>10 days</td>
<td>Receipt of survey materials</td>
<td>Provide feedback on survey materials and methods</td>
<td>ALSWH</td>
</tr>
<tr>
<td>4 weeks</td>
<td>Receipt of survey materials</td>
<td>Finalise survey materials and methods</td>
<td>Collaborator</td>
</tr>
<tr>
<td>4 weeks</td>
<td>Receipt of final materials</td>
<td>Draw sample, final layout, printing, mail surveys or make phone calls</td>
<td>ALSWH</td>
</tr>
<tr>
<td>6 weeks</td>
<td>Mailout</td>
<td>Log returns, conduct reminders, data capture, supply data</td>
<td>ALSWH</td>
</tr>
<tr>
<td>6 months or more</td>
<td>Receipt of data</td>
<td>Analysis and write up results</td>
<td>Collaborator</td>
</tr>
<tr>
<td>Completion</td>
<td></td>
<td>Destroy confidential information, secure storage or destruction as per ethics document</td>
<td>Collaborator</td>
</tr>
</tbody>
</table>
Generally speaking, ALSWH collaborators who have conducted substudies, have managed to maintain adequate progress and to complete their projects in a timely manner. The success of ALSWH substudies has largely depended on clear procedures and timelines at the outset, and ongoing open communication between the ALSWH team and the substudy team.

Costs
Collaborators are encouraged to discuss substudy costs with the ALSWH team before commencement of the substudy to ensure their budget will cover all costs. By using the timeline that has been created, costing from previous projects and quotes from outsourced tasks (eg printing), estimates of resources needed and time required can be made. Costs can then be allocated and a budget created.

Some resources that may be used for the mailout of a paper survey are described in Table 3. For example, a Research Assistant's hourly wage (Rate 2) and the printer's charge per survey (Price 1) are two costs that must be budgeted for. Table 3 shows how the cost of a resource to be used for the mailout of a paper survey is calculated by multiplying the number of units by the unit cost of the resource. The total cost of the mailout is the total of the costs of the resources that are required to complete the task. The total cost of the substudy is equal to the total cost of all tasks that need to be carried out to complete the substudy.

Criteria for selecting substudy participants
As with the budget and timeline, most of the criteria for selecting substudy participants will have been discussed with the longitudinal study team at the time of application. Actually drawing of the sample will follow the longitudinal study's procedures (see Adamson, Young & Byles in this issue) which will take into account the ethical requirements of the study. For example, participants who have taken part in a recent substudy might be excluded from the next substudy, so as to control the extent of participant burden. The following procedures are followed by the ALSWH in drawing substudy samples, and are

<table>
<thead>
<tr>
<th>Table 3: Resource allocation and usage for the mailout of a paper survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of units required</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>N1</td>
</tr>
<tr>
<td>N2</td>
</tr>
<tr>
<td>N3</td>
</tr>
<tr>
<td>N4</td>
</tr>
<tr>
<td>N5</td>
</tr>
<tr>
<td>N6</td>
</tr>
<tr>
<td>N7</td>
</tr>
<tr>
<td>N8</td>
</tr>
<tr>
<td>Total cost of mailout</td>
</tr>
</tbody>
</table>

Notes: This table is used by the collaborator or the longitudinal study to estimate the cost of the resources required for the task of mailing out a paper survey. The costs of other tasks required for a substudy can be calculated in a similar manner and the total cost of all tasks provides an overall cost for the substudy.
offered here as an example of how such procedures might take place.

To draw the substudy sample, the collaborator provides the ALSWH data manager with a clearly defined, written set of selection criteria that detail:

- the subset(s) of the cohort(s) from which the sample will be drawn
- the specific questions and responses from previous surveys upon which the selection is to be made.

For example, participants were selected for an ALSWH substudy on coping with depression by using data for the Center for Epidemiological Studies Depression Scale from the previous two longitudinal surveys (Andresen et al. 1994). Women were identified as depressed or not depressed at each of these surveys. Depression status at the two time points were then used to define four groups of women: those who overcame depression, those who became depressed, those who remained depressed and those who remained non-depressed. The sample was randomly selected so that there were equal numbers of women in each of the four groups.

Before drawing the sample, the ALSWH data manager uses longitudinal study database records to exclude participants based on the ALSWH protocols for selecting substudy participants. Under these protocols, some participants will be excluded, specifically those:

- who do not wish to participate in substudies
- who have already completed a substudy between longitudinal surveys
- who have completed the maximum number of substudies allowed over the life of the longitudinal project
- who are unable to complete the substudy by the chosen research method (e.g. participant may not have a telephone or may have language or eyesight difficulties).

Data and/or participant material that is sent to the collaborator contains only the secondary participant identification number (known as ID alias), thus maintaining the anonymity of the participant (see Graves, Ball & Fraser in this issue for more details).

After final adjustments have been made, including any recent changes to names and addresses, and all substudy criteria have been finalised, the sample can be drawn. (If changes are made during the planning of the substudy, they are dated, defined and documented and the original criteria amended or deleted to avoid any confusion).

3. Substudy administration and documentation

From an administrative point of view, it is helpful to know who has been selected for participation and the status of each participant in a substudy (e.g. whether or not they have responded). This assists staff when participants call with inquiries as they can easily retrieve current information from the database. For documentation purposes, the longitudinal study team needs to know who has participated, in which substudies and when. The ALSWH uses databases to assist with administration and documentation of substudies. An overview of the substudy information contained in ALSWH databases is provided in this section.

Longitudinal study database

A summary database for the longitudinal study stores general information on each substudy, including the:

- substudy identification number
- substudy name
- person responsible for the substudy
- phase number (there may be more than one phase to a substudy, e.g., pilot phase and main phase)
- number of groups. A substudy may have more than one group of participants and each group will have different selection criteria (e.g., the ALSWH depression substudy referred to earlier, had four groups)
- cohort(s) used (e.g., Younger, Mid-age or Older cohort)
• number of participants surveyed
• selection criteria
• significant dates (eg mailout date)
• identification numbers of the selected participants
• whether they completed the substudy.

For record keeping purposes, this database can provide information for each participant including: the number of substudies that a participant has been sent, the number completed and the number and date of their most recent substudy.

Substudy logging database
It is essential to keep track of the participants taking part in the substudy. If collaborators have contact with longitudinal study participants they may collect information that needs to be passed on to the ALSWH team (eg changes of address, participant death). A computerised database is time saving, particularly if the number of participants is large. Therefore, the ALSWH team created a system for logging the return of data see Adamson & Graves in this issue. For example, for a written survey, ALSWH or the collaborator will record the following details in a substudy logging database:
• the date of receipt of the survey
• whether the survey is completed or not completed
• whether the consent form was signed or not
• the receipt of unopened surveys ‘returned to sender’
• participant withdrawal from the substudy and/or the longitudinal study
• participant death details
• changes to name, address or other contact details.

4. Procedure
Once HREC approval has been granted, the survey materials have been prepared, the sample has been drawn and the logging system put in place, the project is ready to proceed. Substudies generally proceed using a framework that is the same or similar to that used by the ALSWH in conducting the longitudinal survey. Not only do substudy materials look the same as ALSWH materials, but the mailing and reminder schedules are also similar - familiarity helps maintain the cohort.

Occasionally, a substudy may elicit phone calls or written correspondence from participants. This correspondence might be a request (eg for health information), or may include other information that requires a response from the ALSWH team either to address a participants’ concerns or as part of the duty of care that researchers must have for study participants. The ALSWH team will respond to these requests as they occur.

Not all substudies use mailed surveys, so the intricacies of substudy procedures will have been determined by the collaborator during the planning stages.

Collaborators sometimes want direct access to longitudinal study participants for substudies and substantial control of the contact and data collection process. ALSWH does not release identifying data without participant consent and ordinarily takes charge of the contact and data collection process in order to protect the privacy of participants and the confidential relationship with participants and to help maintain accurate details for the participant database. Where substudies require direct contact, such as telephone interviews, initial contact and consent to the interview is obtained by ALSWH staff.

As each participant completes the substudy, completion is logged, and data entry procedures commence. Data entry of the substudy may be carried out by the collaborator, the ALSWH team or a data entry company. The same issues that confront the longitudinal study when considering modes of data entry will need to be considered by the collaborator (see Graves, Ball & Fraser; Warner-Smith, Loxton & Brown in this issue).

When data collection is complete, the ALSWH team records all changes to participant details that the collaborator may have received and ensures that the substudy is concluded in such a way that the ALSWH team is aware of participant status.
5. WINDING-UP THE SUBSTUDY
The winding-up procedure for the substudy is just as important as its preparation. Subject to ethical requirements, the collaborator might prepare a written thank you for each of the substudy participants. This could be in the form of a letter or brochure briefly outlining the results of the substudy. Findings of substudies are also included in the annual newsletter to all participants.

The collaborator retains the substudy data set and provides a copy to the ALSWH data manager for archiving. On completion of analyses, the collaborator deletes their copy of any ALSWH datasets and notifies ALSWH, as required under the terms of the MOU. Survey storage and subsequent disposal of substudy materials are undertaken by the collaborator in accordance with ethical requirements.

CONCLUSION
Considerable planning is required to conduct substudies. The success that ALSWH has experienced has been based on several key points. In summary, it is essential to maintain clear communication, to have a transparent application process, to develop a clear framework for procedures, and to maintain documentation. There are advantages and disadvantages in conducting substudies. It would seem however, from the ALSWH experience, that the advantages of conducting substudies outweigh the disadvantages.

Acknowledgements
The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women’s Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

References


A communication strategy can assist a longitudinal study to develop short and long term goals for results dissemination. In addition, researchers are increasingly being asked not just to measure research output in terms of publication of refereed papers, but also in terms of research impact (Cheek et al. 2006). The broader measure of ‘impact’ allows for inclusion of research reports, media contact and research transfer to practising professionals to be included in assessments of research quality. A formal communication strategy that encompasses factors such as the target audience, and types and aims of the various types of communications will go a long way toward capitalising on research. Furthermore, such a strategy can include processes for tracking dissemination, which in turn can be used to assess research output and impact. This paper considers the potential elements of a longitudinal study communication strategy, including developing a distinct public profile, disseminating results to specific audiences, use of a study website and the importance of documenting communication processes and outcomes. The paper draws on our experience in developing communication strategies for the Australian Longitudinal Study on Women’s Health (ALSWH).

Public profile

A key element of a communication strategy is building a public profile for a longitudinal study. Short-term projects are likely to be incorporated into the identity of a particular research centre or institution as part of an ongoing research program. However, longitudinal studies, because of their longevity, tend to have an identity that, while taking account of auspicing bodies, stands apart as an entity unto itself. It can, therefore, be quite important to define the public profile (or corporate image) for a longitudinal study from the beginning of the project.

Consistency in the public profile of the study will be assisted by production of a style manual. A style manual is a document that outlines the
rules for use of the logo, as well as conventions to be used when creating all kinds of media. The aim of a style manual is to create instant recognition by project participants and stakeholders alike – assisting with the longitudinal nature of the project, both in terms of participant retention, funding and research impact.

A graphic designer can produce a comprehensive style manual. While professional graphic designers can be expensive, there are cheaper options that can utilise similar skills. For example, the ALSWH has staff trained to use graphics software so that many elements of the design process can be conducted in-house. At the start of the ALSWH, marketing students contributed to development of the ALSWH public profile as part of their coursework when they undertook a practical placement with the project. It is very important that ownership of the public profile, including images, logos, and all the design elements of the style manual are retained by the study or the institution that auspices the study, and not with the designer. This ensures that use of the logo etc is not controlled by an outside party and will not be subject to extra charges for their use at any point during the life of the project. It is preferable to have editable images on file with the study, so that the study is not reliant on the continuing availability of the designer for production of materials.

Elements of the style manual
The project logo is the cornerstone of a style manual, sets the tone for project colours and style, and is the basis for templates to be used in presentations, reports, and project materials (eg questionnaires). The following considerations might need to be addressed when designing the logo:

- Suitability for a range of media (eg paper, envelopes, electronic files, signage, business cards)
- Applicability to a wide range of purposes (eg surveys, reports, website, correspondence, posters)
- Appropriateness for a variety of audiences (eg participants, funding bodies, academic community, general public)
- Adaptability to both colour and monotone situations
- Compatibility with other logos that might need to be included
- Congruency with applicable trademark laws that may allow you to trademark your logo. Due to national and international differences discussing trademarks is beyond the scope of this paper.

The ALSWH logo was created for the project by a graphic designer at the commencement of the project. In consultation with investigators, the logo was designed to be simple and eye-catching, and incorporated the purple and green colour scheme that was adopted by the women’s liberation/suffrage movement (Sawer 2007). The logo appears on all communications, along with the by-line ‘The Australian Longitudinal Study on Women’s Health’. The logo was designed to suit a range of applications, whether full colour, monotone, large or small format. In addition to the ALSWH logo, all formal correspondence must contain the logos of both universities that auspice the study.

Using the logo as a base, templates for a range of media were then developed and included in the style manual. For example:

- Correspondence (eg letterheads, ‘with compliments’ slips, envelopes)
- Business cards
- Report covers
- Newsletters
- Poster banners
- Powerpoint presentations
- Survey covers
- Web pages.

Having these templates creates a uniform look for all publications, and saves time and expense at the design stage for each product. A style manual could also include guidelines for the use of the...
logo and templates. In addition, any standard statements that the study is required to make when disseminating results should also be included. For example, the ALSWH has approved wording for acknowledging the funding body, and these are included as part of the dissemination of results through any media. Making the style manual available to all people who disseminate project results will assist with creating a public profile and connecting the study to the results being presented.

**TARGET AUDIENCES**

Dissemination of research results is a complex process that involves a variety of audiences. Most commonly, research findings are disseminated through publications in academic journals and in presentations at conferences. In addition, funding bodies may require at least one research report for each funding year. Research reports to funding bodies, including the government, can be distributed to a wide audience of stakeholders and interested parties, beyond the academic community. Results reported in academic journals, at conferences and in official reports sometimes attract media attention, which allows for the dissemination of results to the general public.

**Writing for academic audiences**

The style of writing and presenting for academic audiences is dictated by the discipline and setting (eg journal, type of conference) and communication of results from a longitudinal study does not differ to any large extent from the reporting of any other research design. However, there are a number of points that are important to cover when reporting on longitudinal studies. These points have been considered at great length by a consortium of epidemiologists, statisticians, other researchers, and editors who are developing a checklist for STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) (2007). Some other descriptions of how to report cohort studies are also available in publications by Grimes & Schulz (2002), and Lang & Secic (2006).

Some aspects of particular importance in reporting longitudinal studies are the methods of recruiting and retaining participants. These methods need to be described in sufficient detail to allow the readers to assess the validity and generalisability of the results (Tooth et al. 2005). Loss to follow up also needs to be explained in detail to account for every participant and to assess the risk of bias in results. In addition, longitudinal results are often obtained using complex statistical techniques that may need to be explained in some detail (see Young, Powers & Wheway in this issue for more detail).

The ALSWH has included full descriptions of recruiting and retention methods in published papers (eg Brown et al. 1998; Lee et al. 2005), which can be referred to in later publications. In this way the interested reader can easily source more methodological information. In addition, the ALSWH has produced a set of introductory slides that concisely describe the methods used for recruiting and retention of participants. For example, Figure 1 shows the slide used in most ALSWH presentations that describes the survey schedule. All collaborators have access to these media which can facilitate results dissemination to academic audiences.

**Reports to funding bodies**

Reports to funding bodies are usually required as part of the contract to conduct the research. These reports may vary from single page lists of publications to hundred-page books. Aside from fulfilling funding obligations, dissemination of project findings is a way to maintain financial and academic support for the project by informing stakeholders of the study’s existence and progress. It is necessary not only to describe current results but also to point to future goals, so that the continuing value of the study is apparent to those who are funding or otherwise supporting the study.

The audience for contracted reports can vary from those with technical expertise, such as statisticians, to those who are interested in policy implica-
Consistency in appearance between reports can be maintained by use of the style manual but the style of writing and presentation, and the actual content of the report needs to be appropriate to the intended audience. It is therefore important to liaise with the funding body to establish who the report will be distributed to. In this section we discuss a number of different report types that might be produced by a longitudinal study. Table 1 summarises the three main types of reports that the ALSWH produces on a regular basis.

### Technical reports

By their nature, technical reports are presented in technical language, particularly when discussing statistical or methodological issues. It is appropriate that sections that deal with these subjects would be comprehensive and include sufficient information for a knowledgeable reader to draw conclusions about the methodological and statistical rigour employed by the study. Technical reports can also be a valuable resource for the study and for collaborators, as a historical record of methodological changes and statistical techniques that develop as the study progresses.

### Annual reports

Annual reports are a way of efficiently summarising the study’s progress and should be seen as an important method of communicating the study’s value to stakeholders in particular, and to the community in general. The annual report can also be used as a tool for generating interest in the study, through distribution at conferences and seminars.

![Figure 1: Example of an ALSWH standard slide used at presentations and conferences to describe the study schedule](image-url)

<table>
<thead>
<tr>
<th>Year</th>
<th>Younger</th>
<th>Mid-aged</th>
<th>Older</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td></td>
<td></td>
<td>S1</td>
</tr>
<tr>
<td>98</td>
<td></td>
<td>S2</td>
<td>S2</td>
</tr>
<tr>
<td>99</td>
<td></td>
<td>S3</td>
<td>S3</td>
</tr>
<tr>
<td>00</td>
<td></td>
<td>S4</td>
<td>S4</td>
</tr>
<tr>
<td>01</td>
<td></td>
<td>S5</td>
<td>S5</td>
</tr>
<tr>
<td>02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015+</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
and in providing an overview of the study’s current progress. It can also be valuable to include feature articles on current major projects, abstracts from students who have completed research and higher degrees and abstracts from some (or all) of the papers produced during the year.

**Table 1: Reports to Funding Bodies**

<table>
<thead>
<tr>
<th>Type of Report / Audience</th>
<th>Purpose</th>
<th>Content</th>
<th>Language</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical/Researchers and funding body</td>
<td>• Account of study procedures • Record of study changes • Manual for collaborators • Track study development</td>
<td>Methods and analytical issues such as: • Recruitment methods • Participant retention and response rates • Missing data • Attrition • Data collection (eg survey development and content, interview schedules) • Statistical methods Also included: • Summaries of current research activities • Research team, staff and collaborator information • Study memoranda or meeting minutes • Summaries of output (eg abstracts from publications and presentations)</td>
<td>• Technical • Detailed • Includes detailed tables and graphs</td>
<td>To funding body and to researchers (by request)</td>
</tr>
<tr>
<td>Annual/Broad audience including funding body, government departments and politicians</td>
<td>Communicate study progress and findings</td>
<td>• Welcome letter • Summary of current progress • Details of the research team and study staff • A list of current collaborators and students • A list of research higher degree students who have completed their degrees during the year • A list of all publications and presentations conducted over the year • Any other significant achievements</td>
<td>• Plain language • Written for a naive audience • Highlights more interesting outcomes and activities • Can include simple tables and graphs • Photographs and illustrations</td>
<td>Large distribution list of stakeholders such as politicians, researchers, non-government organisations, media</td>
</tr>
<tr>
<td>Focussed / Funding body, media and other researchers</td>
<td>• Provide answers to a question or issue, informing policy makers</td>
<td>• Executive summary • Literature review • Small amount of information on methodology • Focussed on results, discussion and recommendations</td>
<td>• Focuses on main findings • Emphasise message rather than results and methods</td>
<td>Distributed throughout funding body and to other researchers by request</td>
</tr>
</tbody>
</table>
**Focused reports**

Focused reports that involve major investigation into a particular topic can be a part of the contract with a funding body. Unlike technical and annual reports, these are not used to track study development, progress or achievements, but are undertaken to investigate a research topic to provide answers to a problem or information about an issue, with the main aim of being able to inform policy makers and a related aim of contributing to knowledge.

The target audience of a focused report includes people who are paying for the work to be completed. It can be challenging to determine the extent of the audience of a focused report and to decide how the results of the research can be best presented. Consultations with the funding body can help to meet these challenges. It is also useful to provide a draft report to the funding body, and to allow sufficient time for their feedback to be incorporated into the final version of the report.

Report writing is different to writing for an academic audience, although the components are similar. For example, both will include an introduction, some details about the methods and statistical analyses that were used, the results, and some sort of conclusion or discussion. However, the order and style of presentation, and the type of information included in these sections can be quite dissimilar.

In introducing a report, the audience is interested in past research but is also likely to be interested in where the research sits in relation to policy. For example, when discussing health service use in rural areas, it could be appropriate to include a summary of recent government initiatives in that area. Unlike an academic paper, it would be unusual to focus heavily on past theoretical developments in the research area and more usual to focus on known outcomes. For instance, in a report on the health of women who have experienced abuse, focussing on past theories of why abuse occurs would be less relevant than reporting on past research that has shown associations between abuse and health. In summary, where an academic paper may justify research questions in terms of past research and theory, a report will justify research questions in terms of current knowledge and the need for specific information to inform policy development.

While it is important to include details of methods and analyses in a focussed report, inclusion of such details can detract from the main message of the report, which should concentrate on major findings. It can therefore be useful to include methodological issues, such as statistical techniques, study background and method in appendixes or a separate technical report. This is a quite different approach than that used in academic writing, where the methods and analyses precede the research findings.

Arguably, the most important section of a focussed report is the Executive Summary. This is the section that most people will read, and content here may direct people to continue on to specific sections in the body of the report for further details. This section needs to highlight the major findings in a way that is both brief and accurate. While it may be necessary to flag some caveats that are discussed further in the report, this is generally not the place for speculative findings or for debate. Rather, the summary should focus on those findings that are firm, credible and meaningful in the context of the study.

Although presented first, the Executive Summary will usually be written last. It is likely that this summary will be more widely distributed than the entire report, and so it needs to be written as a stand alone document. As the major conduit through which the results will be disseminated, the main messages of the results need to be concisely and clearly included with a minimum of qualification. Unlike an academic abstract, the executive summary can be relatively long, and is focussed on results and policy implications. Furthermore, the language used must be targeted for a diverse audience, including the general public.

The next most important section of a focussed report is the results section. It is here that the
interested reader will centre their attention, with a view to using the information to identify problems, and to develop solutions to those problems. The results need to be clearly presented and information that will potentially obscure the findings (such as a discussion on the relative merits of alpha being set at 0.01 or 0.05) should be placed elsewhere in the report. Graphs and tables are useful because they offer a quick summary of the findings and can be easily copied from the report to other documents. For example, a graph that shows an increase in obesity over ten years can easily be added to a ministerial brief. Similarly, sentence structure needs to be carefully considered. It is possible that the policy maker may want to use quotes from the report to support a recommendation, for example, a government committee. Brief ‘take away’ messages such as ‘one in four women experience domestic violence’ can be easily visualised and recalled. Also, proportions such as ‘one in four’ have more inherent meaning for non-statisticians than a percentage estimate such as 25%. It is possible that parts of the report will be used without the context of the entire report. It is, therefore, very important that information segments can stand alone and be understood.

The audience for a focussed report will want to know what the research findings imply for policy, and might even have requested that the researchers make specific recommendations in this regard. This requirement should be met in the final section (eg discussion, conclusion) of a focussed report. While elements of academic discussions are important to include (eg where the research fits in relation to past research, what it adds to knowledge, the limitations of the methods and to directions for future research) they are not as important as the policy implications of the research findings. It is also important to bear in mind the organisational constraints that may apply to policy recommendations. For instance, while people working in a government health department may appreciate the contribution of general education to public health, they may not be in a position to influence this arena of activity. If this is the only policy relevant finding in your report, the health department may have difficulty making use of it.

As the preceding discussion has stressed, determining the target audience is perhaps the key element in determining how research results will be disseminated, including the style of writing and the presentation of results. By way of an example, Table 2 presents the same findings for different audiences. The findings relate to an ALSWH focussed report that had examined risk factors for chronic conditions (Brown et al. 2006). The focussed report results were detailed, included some technical language, and referred to tables that contained more detail. The technical details concerned with methods and results were contained in appendixes. The executive summary of this report condensed these results into a much shorter statement, while retaining the policy relevance. The annual report also included a summary of the findings, but all technical language was removed. Media coverage, discussed in the following section, is another useful research transfer method, which also requires attention to the target audience. In Table 2, the heading and first paragraph of the press release for the ALSWH report reduced the findings to a ‘sound bite’ designed to attract attention while not being sensationalist.

The Media

The general public will mainly become aware of the study by media coverage. This can occur as a result of research findings being published or presented, or as a result of a focussed report being published and released by the government. The media is an important means of maintaining a high public profile which is helpful for funding the project. In addition, results disseminated by the media is a useful research transfer strategy (Woloshin & Schwartz 2004). However, there can be tension between media releases and publications – as research papers are usually governed by the originality requirements of the scientific journals they are submitted to.
TABLE 2: CONTRASTING PRESENTATIONS OF RESULTS FROM THE AUSTRALIAN LONGITUDINAL STUDY ON WOMEN’S HEALTH (BROWN ET AL. 2006)

<table>
<thead>
<tr>
<th>Focussed report</th>
<th>Executive summary</th>
<th>Annual report</th>
<th>Press release</th>
</tr>
</thead>
<tbody>
<tr>
<td>There were very strong trends in prevalence of hypertension with increasing body mass index for all three age cohorts (with prevalence approximately doubling across categories in the Younger women). Incidence of hypertension also increased markedly with body mass index in the Younger and Mid-aged women (but not in the Older women) (Table A 3.2)</td>
<td>The most striking feature of the results presented in this report is the adverse effect of overweight and obesity on the prevalence and incidence of vascular disease (hypertension, heart disease and diabetes) as well as asthma. In comparison none of the other risk factors examined showed such consistent and strong associations with chronic conditions. In light of the increasing weight in all age groups, weight gain clearly poses a major threat to the health of Australian women.</td>
<td>Being overweight is the most important risk factor for chronic disease in women at all ages. Women who are overweight are more likely to have hypertension, heart disease, diabetes, osteoporosis and asthma. In comparison none of the other risk factors showed such consistent and strong associations with chronic conditions.</td>
<td>Maintaining healthy weight – the key to avoiding chronic disease Maintaining a healthy weight has emerged as the most important factor for Australian women in avoiding chronic diseases such as vascular disease, diabetes and asthma, according to researchers from two Australian universities.</td>
</tr>
<tr>
<td>For both the Mid-aged and Older cohorts there were very strong and statistically significant trends with prevalence and incidence of reported heart disease increasing with body mass index (Table A 3.8).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The increases in both prevalence and incidence of diabetes with increasing body mass index were striking and statistically very significant in all three age cohorts. These associations were among the strongest in this section of the Report and provide unambiguous evidence of the graded risk of diabetes associated with weight (Table A 3.14).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevalence and incidence of asthma increased significantly with increasing body mass index for all three age cohorts of women. The consistency of this pattern points to an important, potentially causative relationship (Table A 3.20).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body mass index is usually found to be inversely associated with osteoporosis. In these data, however this trend is clear only for prevalence among the Older women. While other differences between groups defined by body mass index are statistically significant they are less consistent and could be affected by biases associated with diagnosis (Table A 3.26).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevalence of arthritis at Survey 1 increased significantly with increasing body mass index for both the Mid-aged and Older women. This pattern was not apparent however for incidence in either cohort, which is difficult to explain (unless women gained weight after a diagnosis of arthritis, possibly due to difficulty taking sufficient exercise) (Table A 3.32).</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The main point of contact between a longitudinal study and the media is a media release, written by a researcher and distributed usually through their institution to the major news agencies. Wherever possible, a coordinated media release in conjunction with the publication of a major report, paper or presentation will enhance the impact of the research findings and raise the profile of the study.

Most academic institutions have media offices that can assist in the preparation of a media release, and should be the first contact point for the media. It is important to consider who will respond to requests for interviews and, if possible, to arrange media training for interviewees prior to contact with the media.

In responding to media requests, there are a number of things to bear in mind. It pays to have a simple message that can form the basis of ‘the story’. Is there a ‘health promotion message’, or ‘a warning’, or do you want to raise awareness, challenge prevailing attitudes, or report ‘good news’?
Don’t expect that the media will necessarily have grasped what is new and interesting about the findings, nor the context from which they come. In a large study such as the ALSWH that covers many issues, there is a danger that the impact of any one finding may be lost in the message about the study as a whole. Equally, the study does not need to be associated with a single issue. Finding a balance between the current message and the overall aim of the study is important.

It is also important to pitch the message at an appropriate level. The findings may be sensational, but you don’t want them to be sensationalised. At the ALSWH, we feel an obligation to our participants to ensure that the information they provide to us is represented fairly and is not ‘used against them’. For example, findings about the high incidence and prevalence of obesity among Australian women had tremendous public health significance and contributed to an important health promotion goal. However, there is a danger of the data being used to accuse women of being solely responsible for their own ill health because of their weight. Sometimes it is difficult to convey more complex and subtle information, such as the role of socioeconomic and other environmental factors that contribute to an individual’s behaviour.

Communication with media is different from communicating with the scientific community. In scientific writing we have the opportunity to place findings in context and to argue their strengths and weaknesses and whether they are consistent with other data. In communicating with the media there is no room for such complexity or debate, and the message must be kept clear but not misleading. Our message about safe levels of alcohol intake for older women (Byles et al. 2006), for example, might easily have been translated into a message that ‘alcohol is a path to healthy ageing’. However, this was not the aim of the study or our interpretation of the findings. We had to be clear that our message was about whether there was any evidence that it was safe for women aged over 70 to drink two standard drinks each day.

Also, in scientific writing we tend to objectify the participants, but this is not appropriate when dealing with the media. In talking to the media, you are potentially talking directly to your participants. One trick is to imagine that the reporter is one of the participants in the study. In this way you will always pitch your message at an appropriate level and be respectful of study participants.

Another issue in relation to media reports is whether participants should be contacted for interview. We hold that participants’ details should never be provided to media even if they are willing to be interviewed. Our ethical undertaking is that no identifiable individual data will be reported. Even if the participant concerned is comfortable being identified, other people in the study may not be comfortable and may come to doubt the assurance of confidentiality.

Know the media you are talking to and what their interest will be – for example a broadsheet newspaper will have a different audience to a science-based radio program. It is also important to understand, if possible, any issues the interviewer may want to raise. For example, an investigator may be taken off-guard during a hostile live-to-air interview, when the agenda of the interviewer may clearly outweigh the findings being reported. Tactics for dealing with such unexpected situations can be learned in media training courses, and this can be of use to investigators from high profile studies. And one related and final point: never assume that anything is ‘off the record’.

Documenting communication processes and outcomes
It is useful to maintain one central record of all publications, presentations, reports and media relating to the study. Further details on how this record is maintained are provided in Chojenta, Mooney & Warner-Smith in this issue. This is an important measure of the impact of the study. It also helps in developing an overall profile of the study and its contribution to information, debate and policy.
WEBSITES
A comprehensive website can be a focal point for communication for a longitudinal project. A website is the first port of call for many collaborators, funding body members and other stakeholders for information regarding methodology, current progress and publications. In addition, participants, service providers, and the general public are likely to access the study website. Thus, the website has a wide audience and information needs to be appropriately pitched. Resources should be allocated towards not only the establishment of the website, but also to the maintenance and development of the site over time.

When establishing a website it might be useful to consider:
• Including a space for participants to view information and study results
• Including a space for collaborators and investigators with a password to protect sensitive information
• Contracting a website designer to establish the site
• Asking the designer to create the site in such a way that it can be edited by the project staff, at least for minor changes
• Who will host the website? The study’s institution may offer adequate website hosting or there might be a need to search for an external website host
• Purchasing the rights to a customised domain name. This can be an effective means of placing your site. Be mindful to purchase a name that will be easily remembered and replicated – the shorter your address is, the better.

The ALSWH project website has developed into a comprehensive source of information on the project, and is continually updated and developed. It includes an outline of the project methods and sample, and details of all project analyses and sub-studies. It also contains electronic copies of many of our documents and outputs. The ALSWH website has been designed to communicate information to several different groups including participants, new and existing collaborators, and staff and investigators. A section dedicated to disseminating information to participants was recently upgraded to include summaries of results for each cohort. A comprehensive archive of information for collaborators can be accessed on the project website, including copies of surveys, data books, policy documents and data documents. The project website has become an invaluable tool for communicating information to staff and investigators across the two main sites of the ALSWH. The website allows staff to access minutes of meetings, and other internal documents from a secure password-protected site. Maintenance of the ALSWH website is managed internally, and all substantial changes to design and content are approved by the ALSWH Steering Committee. For more detail, the ALSWH website can be viewed at www.alswh.org.au.

MEASURING IMPACT
Measuring the impact of media, reports and other outputs is a complex task for any researcher. However increasing importance has been placed on attempting to quantify the impact and use of research outputs. Impact may include not only recognition by publication in prestigious journals, but also what use is made of the research finding, the influence on policy and practice and the potential for these changes to benefit health. The ALSWH research team are currently investigating effective methods of measurement of these impacts, and will continue to develop methods in this field.

CONCLUSION
Through the development of a communication strategy, the ALSWH have not only been able to effectively disseminate research outputs to the intended audiences, but also track with a view to measuring the impact of these outputs. Through careful planning of the Study’s public profile, the ALSWH has developed a clear and recognisable image not only within the academic realm, but also to a wider audience. The Study’s image has been established with a recognisable logo and corporate style, and importantly has an online presence. The website provides participants with a
means of communicating with researchers, and is a focal point of communication with staff and new and existing researchers at other centres.

Acknowledgements
The research drawn on for this paper was conducted as part of the Australian Longitudinal Study on Women’s Health, The University of Newcastle and The University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding and to the women who provided the survey data.

References
Tooth L, Ware R, Bain C, Purdie D and Dobson A (2005) Quality of reporting of observational longitudinal research, American Journal of Epidemiology, 161(3): 280-288; response to letters to the editor, 2005 162: 1033

Free inspection copies of International Journal of Multiple Research Approaches are now available for consideration by your Faculty’s library selection committee. Please photocopy the form below and forward it to the committee representative in your Faculty. We shall send you and evaluation copy on request – please contact the Journal promptly at: subscriptions@e-contentmanagement.com

I recommend that the library subscribes to International Journal of Multiple Research Approaches (ISSN 1834-0806 - sample material on www.ijmra.com) for the following reasons:
☐ I have evaluated the content of this journal and can confirm that the high quality of the articles will be of significant benefit to faculty and students in teaching and research
☐ The journal has core content which is central to the research and teaching interests of my department
☐ I am an author / member of the editorial board and as such will be using it regularly for research, article submission and as a teaching aid.
☐ I will be using this journal for my student course packs and reading lists
☐ I will be using this journal for my own research and scholarship
☐ I have noted the journal is frequently cited and is therefore a critical acquisition for our collection in this subject area

Requested by: ____________________ Faculty: ____________________ School: ____________________
Signature: ____________________ Date: ____________________ Email: ____________________
ORDER FORM

CONDUCTING LONGITUDINAL RESEARCH:
PRACTICAL LESSONS FROM THE AUSTRALIAN LONGITUDINAL STUDY ON WOMEN’S HEALTH

Edited by
Deborah Loxton, Project Manager, Australian Longitudinal Study on Women's Health, University of Newcastle
Julie Ellen Byles, Director, Research Centre for Gender, Health & Ageing, University of Newcastle
Annette Dobson, Professor of Biostatistics, School of Population Health, University of Queensland
Wendy J. Brown, Professor of Physical Activity and Health, School of Human Movement Studies, University of Queensland

ISBN 978-0-9775742-9-2; ii + 134 pages; softcover

A special issue of the International Journal of Multiple Research Approaches
(Volume 1, Issue 2 – ISSN 1834-0806, December 2007)

North America, Europe, North East Asia & Middle East:
Institutions US$121.00; Individuals US$73.50, incl. airmail
Australia, New Zealand, Pacific, China, South & South East Asia, South America & Africa:
Institutions AU$121.00; Individuals AU$73.50, incl. GST and airmail

This special issue of the International Journal of Multiple Research Approaches is also available as a course reader; course coordinators are welcome to request an evaluation copy. Students receive a 20% discount on rrp.

Number of copies: Total: US$ Total: AU$

A cheque made payable to eContent Management Pty Ltd is enclosed OR

Please charge this purchase to my Mastercard/Visa (a 4% merchant fee applies):

My full card number is: _____ / _____ / _____ / _____ Expiry date: _____ / _____

Cardholder Name: ___________________________ Signature: ______________________________

Please send this order to the address below by airmail:

Name of Institution: _________________________ Name of Individual: _______________________

Street address (or PO Box):

City: ______________ State: ______________ Zipcode: ______ Country: _____________________

Tel. (with area code): ______________________ Fax (with area code): ______________________

Email: ___________________________ Your reference: ______________________________

Fax the completed order form to +61-7-5435-2911 or post with cheque to:
eContent Management Pty Ltd, PO Box 1027, Maleny QLD 4552, Australia
Tel.: +61-7-5435-2900; Email: subscriptions@e-contentmanagement.com
www.e-contentmanagement.com