Theoretical sampling designs for a UK birth cohort with potential accelerated design

Paul A. Smith*
James Dawber*
Peter G.M. van der Heijden*†

*S3RI and Dept of Social Statistics & Demography, University of Southampton
†Department of Methods and Statistics, Utrecht University

July 2019

This work was funded by the Economic & Social Research Council under grant no. ES/T001038/1.

Theoretical sampling designs for a UK birth cohort with potential accelerated design

Paul A. Smith*, James Dawber*, Peter G.M. van der Heijden*†
*S3RI/Dept of Social Statistics & Demography, University of Southampton, UK
†Department of Methods and Statistics, Utrecht University, the Netherlands

Executive Summary

Accelerated Longitudinal Designs (ALDs) have several advantages over single cohort designs, specifically the ability to estimate and adjust for cohort effect differences and produce results in a shorter period. These are traded off against additional model complexity in analysis and challenges in recruitment of additional sample cases in the field.

There are several properties of ALDs which can be varied; analyses based on the power to detect cohort differences in a linear model for age effects show that small numbers of cohorts have the best chance to improve the properties of the design and speed up data availability across the age range of interest while remaining practical to implement.

Further work is needed on methods to determine sample sizes to meet variance constraints for a range of analytical parameters, and to determine how best to produce a general purpose design which can meet a wide range of user needs.

1 Introduction

When studying change there are two broad designs which can be used; cross-sectional and longitudinal. The former design compares different age groups at a single point in time, whereas the longitudinal design measures subjects more than once at different time points. Often longitudinal designs follow one cohort over time, encompassing multiple measurements. The benefit of such a design, as opposed to a cross-sectional design, is that variations *within* individuals can be observed, not just *between* individuals. This within-individual variation over time offers greater insight into the change being studied, and provides the opportunity to make causal statements. However, the obvious disadvantage is that the length of the study must be long enough to capture this change, which in some cases can be many years long. With this longer study time comes cost and logistical challenges, as well as higher participation burden contributing to higher rates of dropout (Moerbeek, 2011). For example, suppose dropout was 10% per wave and there were ten waves in the longitudinal study then only 35% of the original sample would remain at the end. This creates limitations in the generalisability of the research as the remaining sample is likely to have only small numbers of cases in some subgroups of interest and reduce representativeness.

An accelerated longitudinal design (ALD) provides a way to overcome some of these challenges. An ALD was first introduced by Bell (1953), and such a design uses multiple cohorts of differing ages which collectively capture the entire age range of interest. Consider again the above example with a longitudinal design with ten waves and a dropout of 10%. An ALD could be designed with two cohorts each covering five waves, not only would this design halve the study length but also 59% of the original sample would remain, compared to 35% for a single cohort. This highlights the primary advantages of an ALD; a reduction in the study length and with it lower total dropout. In this report we define an ALD as any longitudinal design which utilises more than a single cohort. These cohorts can either by sequential or overlapping, as summarised in Figure 1. The example in Figure 1 presents three longitudinal designs which cover ages 0-8 years with three cohorts for the two ALDs. Note that the total design times in this example are nine, three and five years respectively, highlighting the potential reduction in time for accelerated designs. An ALD has also been referred to as a cross-sequential design (Schaie and Strother, 1968). We further discuss the advantages as well as disadvantage to an ALD in Section 2.

A further exploration of ALDs is presented in the next section, focussing on the advantages and disadvantages. Following this we introduce how data from ALD designs are modelled, so that inferences can be made. Next, methods to optimise the design of an ALD are discussed, in regard to number of cohorts, extent of overlapping and frequency of measurement, followed by tests of convergence. The next section explores the pros and cons of flexible age design as opposed to fixed age design, and lastly examples of ALD are introduced with commentary on their performance, and a short evaluation of possible designs for a new cohort is made.

Figure 1 Examples of an accelerated longitudinal design compared to a single-cohort design

Single-cohort longitudinal design									
Cohort 1	○	○	○	○	○	○	○	○	○
Accelerated longitudinal design - Sequential									
Cohort 1	○	○	○						
Cohort 2				○	○	○			
Cohort 3							○	○	○
Accelerated longitudinal design - Overlapping									
Cohort 1	○	○	○	○	○				
Cohort 2			○	○	○	○	○		
Cohort 3					○	○	○	○	○
Age	0	1	2	3	4	5	6	7	8

2 Accelerated longitudinal designs

2.1 Advantages

As discussed, one of the primary advantages of ALDs compared to single-cohort designs is its ability to reduce the duration of a longitudinal study while still covering a relatively large age range. This reduction in duration has many secondary advantages. Firstly, the funding may be more certain over a shorter period, so that the survey objectives can be met within a defined funding envelope without the need for repeated renewal of funding (although this does not preclude continuing the study if further funding is available, of course). Secondly, since the subjects require fewer measurements the burden per participant is reduced, leading to lower attrition (over the timescale of the ALD). Keeping attrition low is very important to ensure that the original sample remains representative. A further advantage to an ALD is a decrease in potential retest effects, for example subjects are less likely to provide misinformation because they have learnt that that leads to reduced interview times. Finally, the shorter period means that the concepts and methods for data collection are less likely to change during the course of the survey, so models and comparisons are less likely to be affected by evolving survey practices and objectives.

Another significant advantage to an ALD is an increase in generalisability. A single-cohort design is naturally dependent on the targeted cohort which may not be generalisable to the wider population. For example if the cohort is all children born in a given year, the findings are less likely to be generalisable to all children born in future years. This is because the study findings may be due to a cohort effect. A common challenge with longitudinal studies is the confounding of this cohort effect with age and period effects. A cohort effect is a change relating to people born at a given point in time, independent of the effects of ageing. Age effects are changes which occur as people age regardless of which cohort or period they are in. And lastly, a period effect is a change occurring at a given point in time which affects people of all ages and all cohorts. It is useful for the age effects to be identified because these effects are generalisable to the wider population. An ALD covers multiple cohorts, which means that there is some information to analyse the cohort effect (it is not confounded with other changes), leading to more generalisable findings. In particular it is possible to test whether there is a difference in the cohort effects. However, if a cohort effect is present it can lead to greater challenges in analysis and inference for the combined data from the ALD, as does controlling for period effects.

2.2 Disadvantages

An important disadvantage of ALDs is the possibility of a cohort effect. If this arises then there are inherent differences between the individuals in the cohorts, which makes it challenging to develop a model for age effects which might be generalisable to the wider population. Methods which test for such a cohort effect have been explored, often using random effects models (see for example Miyazaki & Raudenbusch 2000, Galbraith *et al*. 2017). Such methods that examine the overall trajectory over the age range across cohorts are referred to as tests of convergence. A convergent design is one that does not have a significant cohort effect, meaning the age effect can be measured. These tests of convergence are explored in Section 5.

While single-cohort designs require many design considerations such as frequency and timing of measurements, ALDs require the added considerations of the number of cohorts and the extent to which the cohorts overlap (if at all). These additional factors create added complexity in sample size and power calculations. These factors and how they should best be used in practice are presented in Section 4.

The interest of the study may be specifically for a certain cohort, for example the Millennium Cohort Study specifically focussing on people born at the start of the new millennium. Although not a disadvantage of ALD, it would offer no benefits to the study if the focus is on measuring a specific cohort. The UK has a strong history of cohort studies, which have provided opportunities to do cross-cohort research. Such research is implicitly doing a multi-cohort analysis, similar to such an analysis through an ALD. One could argue that the intentions of setting up national cohort studies approximately every ten years is implicitly an ALD without explicitly designing them this way.

3 Models for ALD

Longitudinal studies are designed to measure change, and in order to do this a suitable model must be used. This model must incorporate the correlation within each subject, usually with a hierarchical model with measurements nested within subjects. This hierarchical design is easily adapted to ALD by simply adding another level to the model, with subjects nested within cohorts. Often multilevel linear regression models are used for this hierarchical framework, and these may also be called hierarchical linear models and random-effects models. Structural equation models (SEM) can also be used to model ALDs, but they have been shown to be merely a more general form of a multilevel model (Curran, 2003) and the similarities often outweigh the differences (Curran *et al*., 2010).

The ideal model will naturally depend on the specific requirements of the study. Some considerations include the type of response variable being measured, which covariates to include as fixed or random effects, what longitudinal trend is expected, i.e. linear or polynomial, and what interactions between covariates are expected? One example of a general multilevel model for an ALD is presented by Galbraith *et al*. (2017):

$$y_{ij} = X_{ij}\beta + W_j\gamma_j + Z_{ij}b_{ij} + \varepsilon_{ij} \tag{1}$$

where y_{ij} is the vector of responses for subject i in cohort j, X_{ij} and Z_{ij} are design matrices for the fixed and random effects, W_j is the random effects design matrix for cohort j, β is the parameter vector for the fixed effects, γ_j and b_{ij} are the vectors of the random effects for the cohort and subject levels respectively and ε_{ij} is the residual vector for each subject.

Once a specific model is constructed it is then possible to construct sample size calculations, calculate the statistical power and statistical inferences that will be required. One such inference is a test of convergence which will be discussed in the Section 5.

4 Optimal design

4.1 Single cohort designs

For a single-cohort design a number of factors need to be considered including the study duration, the number of subjects, and the frequency of measurements. Raudenbush & Liu (2001) analyse the effects of these factors on the power to detect group differences in a single-cohort design. They conclude that increases in sample size increase power, as would be expected. More importantly, they show that increasing frequency and duration generally increase power though this is dependent on what longitudinal trend is being modelled, as well as the variation between and within subjects. Moerbeek (2008) shows that with no dropout the study duration has a stronger effect on power than number of measurements, which is heightened if there is a higher order polynomial trend. However if there is dropout, increasing the study duration can have a negative effect on power. Ultimately the power and hence the optimal study design depend on associations between many factors which cannot be controlled in the design or modelled easily, such as dropout rates and patterns, variance components, missed measurements, and degree of polynomial trend. Both Raudenbush & Liu (2001) and Moerbeek (2008) present frameworks for dealing with this complication.

4.2 Accelerated longitudinal designs

While a single-cohort design has its own challenges, an ALD has two additional considerations: the number of cohorts and the extent of overlap between the cohorts. Despite this added complexity, Moerbeek (2011) and Galbraith *et al*. (2017) discuss in some depth a framework for designing an accelerated longitudinal study. If attrition can be assumed to be absent from a study then Moerbeek (2011) suggests that two cohorts are preferable to one or three, and also that the length of overlap of one year is a good compromise. In this case a high overlap increases the power when the number of subjects is fixed, but decreases power when the number of measurements is fixed.

A survey with no attrition is, however, very unlikely in practice, so the decrease in cohort sample size over time must be taken into account. Naturally if attrition occurs then shorter cohorts are preferable to reduce its impact, especially when attrition is large and occurs early in the study (Moerbeek, 2011). Galbraith & Marschner (2002) postulate a Weibull-based model for attrition which fits a parameter showing whether attrition is constant, or concentrated earlier or later in the study. Galbraith et al. (2017) find that this model is a good fit for the attrition pattern in the Longitudinal Study of Australian Children, which indeed shows attrition to be higher earlier in the study. Using shorter cohorts will require more cohorts to cover the age range and reduce the attrition; but keeping the frequency of measurement low should also be considered, as high frequencies may be the direct cause of the attrition.

Galbraith *et al*. (2017) undertake an evaluation based on a linear model for a survey outcome with age (a linear trend for an age effect). Under this model they provide an expression of the power, and use this to compare different designs. They show that to detect a linear trend with fixed power, the number of measurements per subject (rather than number of cohorts and extent of overlap) is the primary determinant if either the number of measurements or subjects is to be minimised. If the number of measurements per subject is fixed, then a design with fewer cohorts is preferred, giving less overlap. If a linear trend is being modelled then the frequency of measurement has little effect on the required number of subjects, suggesting that the number of subjects is more important than the number of measurements. It should be noted that these findings were based on models assumed to be correct, and this assumption cannot be tested in practice. Furthermore, it assumes a single response variable whereas in practice there are multiple outcomes that need to be considered. This further complicates identifying the optimal design for an accelerated longitudinal study.

5 Tests of convergence

Harring et al. (2016) suggest that after applying an ALD, the first step should be to test for convergence. This is important as it can test whether the cohort effect of the ALD can be assumed to be ignorable or not, and if so the age effect can be measured. If there is a significant cohort effect then this can distort trends over age due to divergent trends between cohorts. However if the test shows the cohort effect to be negligible then the multiple cohorts can be assumed convergent and inference on age trends can be made reliably, so long as period effects are assumed ignorable.

Convergence was first tested by Bell (1954) using significance tests of the means where two cohorts overlap at a certain age. However two types of methods have developed which more rigorously test the cohort effect. These methods test the convergence directly in the model structure, rather than through a post-hoc approach. Also, which test of convergence is used is dependent on the structure of the ALD model and also whether a SEM or multilevel model is used for estimation.

Using an SEM framework allows a latent curve growth model to be used which can be used to test for cohort effects. In this framework parameters of individual change are defined as latent variables or factors. By viewing cohorts as subpopulations the SEM can be constructed to test for differences between cohorts. Duncan et al. (1996) apply this method to test for cohort effects in a study looking at changes in alcohol use. The SEM approach is highly flexible when it comes to model specification (Wu *et al*. 2009), but it is most easily applied in a situation with the same sequencing of data collection across participants ("balanced in time" (Miyazaki & Raudenbush 2000), possibly with missing data as long as it is missing at random). It is possible to fit SEMs where the patterns of observation are not consistent using full information maximum likelihood, but not to calculate the same fit indices to assess the convergence (Wu *et al*. 2009).

Miyazaki & Raudenbush (2000) presented a method using multilevel models in an example measuring antisocial attitude changes between subjects between 11 and 21 years of age. This method requires two models to be fitted, the first which is able to take cohort differences into account and the second that does not. The fit of these two models are then formally assessed using a likelihood ratio test or Wald test. If there are significant differences then this suggests there is a cohort effect. A benefit of this framework is that the ages do not need to be fixed, as the model can account for flexible ages. A further benefit is that it is relatively simple to incorporate a further hierarchical structure into the model such as a cluster like a family or a school.

If a cohort effect is detected, then a more complicated analysis may be required, however anticipating potential cohort effects in the model design enables them to be accounted for. It is important to first discern whether the cohort effects can be ignored or not, hence the importance of these tests of convergence. If a cohort effect is expected then a single-cohort design should be preferred for analysis of age effects (Moerbeek, 2008, Duncan et al., 2006), although it will not be generalizable to other cohorts.

There may be other ways to test for convergence. Buscot et al. (2017) propose a Bayesian model for detecting divergence between longitudinal trajectories, and it seems possible to adapt this approach to provide a test of convergence.

6 Fixed vs flexible age design

A longitudinal design can either have a fixed or flexible age design. A fixed design is when the repeated measurements are made at specific ages for each subject, while a flexible design allows for more variation in the age of the subject when measurements are made.

The advantages of a fixed-age design are mainly logistical, since questionnaires and tasks only need to be designed for a single age group. Methodologically there are no clear advantages of a fixed–age design, except for the effects on measuring convergence using SEMs as described in section 6.

The main disadvantage of the fixed age design is that it results in large age gaps between measurements from which no conclusions can be drawn. For example, if measurements are made at ages 10 and 15 it is impossible to know whether the trend between these two ages is linear or not. However this problem is less significant if numerous measurements are made over time, where the trends should be more identifiable, particularly if linear.

The primary benefit of the flexible-age design is that it provides more information about the trajectory curves over time. This design will lead to each cohort having a range of measurements across the desired range of ages. This design is statistically feasible, as multilevel models can deal with this kind of information.

A possible disadvantage of having variable measurements is that a specific function (variable) cannot always be measured with the same instrument at each age point. However, this is also true when measuring more frequently in different subgroups. In longitudinal studies that cover a large age range, functions may be measured at different ages using different instruments. Performance on that function over time will therefore have to be normalized to z-scores or ranks to infer developmental trajectories. An example of this is the measurement of IQ which requires different instruments at different ages. The IQ score is naturally normalised to ensure comparability over time.

7 Examples of studies using accelerated longitudinal designs

A number of large longitudinal studies have utilised an ALD, in this section we briefly introduce some notable ones from around the world.

7.1.1 Longitudinal Study of Australian Children (LSAC) – Australia

The LSAC study uses two cohorts, but with a large overlapping period (Gray and Sanson, 2005). The first cohort represented Australian infants aged 0-1 years, while the second cohort represented Australian children aged 4-5 years old. In total there are eight waves taken every two years, the final wave occurred in 2018 with the ages of the children being 14-15 and 18-19 years in respective cohorts. The Australian Longitudinal Study of Indigenous Children (LSIC) runs alongside the LSAC and utilises a similar accelerated design with two cohorts (Thurber et al., 2015).

This specific ALD was chosen as a means to make the results more generalisable to the wider population as opposed to one single cohort. With such a large overlap the ALD was clearly not chosen to reduce the study duration. Further to the ALD, the study takes on a cluster design where 330 postcodes were random selected and children were selected from these postcodes. A stratification was also applied to ensure the sample was proportional in terms of state/territory as well as urban versus non-urban.

7.1.2 The National Longitudinal Survey of Children and Youth (NLSCY) – Canada

The NLSCY is a longitudinal study with multiple cohorts but primarily to boost the younger ages (Michaud, 2001). An original nationally representative sample of 0-11 year olds were randomly selected and as this cohort aged, new cohorts of 0-1 year olds were added to the study. This ensured that nationally representative estimates could continue to be made for younger ages.

7.1.3 Growing Up in Ireland (GUI) – Ireland

The GUI study began in 2006 and used a two cohort design with one cohort for infants (9 months old) and the other for children aged 9 years old. Both cohorts were representative random samples of 11,000 and 8,500 respectively. The study is currently in Phase 2 where the first cohort are approaching 13 years old and the second cohort approaching 20 years old. This means the study is 13 years old and there is currently an overlap of four years over an age range of 0-20 years.

7.1.4 Growing Up in Scotland (GUS) – Scotland

GUS incorporates three cohorts in the design. When the study began in 2005 two cohorts were selected from randomly selected areas: a 0-1 year and a 2-3 year old group. Both these cohorts were kept in the study until they were six years of age, however data from the younger cohort was continued until the start of secondary school. A third cohort of 0-1 year olds was added in 2011.

7.1.5 The Swiss Survey on Children and Youth (COCON) – Switzerland

The COCON Study covers three randomly selected cohorts: middle childhood (6 years), middle adolescence (15 years) and young adulthood (21 years) over a period of 15 years (2006-2021). The age range of interest is between 6-21 years so the young adult cohort actually are only measured once (Buchmann, 2017).

7.1.6 YOUth Cohort Study - Ultrecht

The YOUth Study began in 2015 and uses two cohorts each with a sample size of 6000, including a 'Baby & Child' cohort capturing the age range of 0-8 years and a 'Child & Teenager' cohort capturing the ages of 9-15 years. There is planned overlap between the two cohorts at ages 8-10, although the funding for this is currently in doubt, and a flexible-age design is used with approximately a three year range for each wave. The YOUth Study is not a national study, focussing just on Utrecht and its surrounding areas.

7.1.7 PAIRFAM – Germany

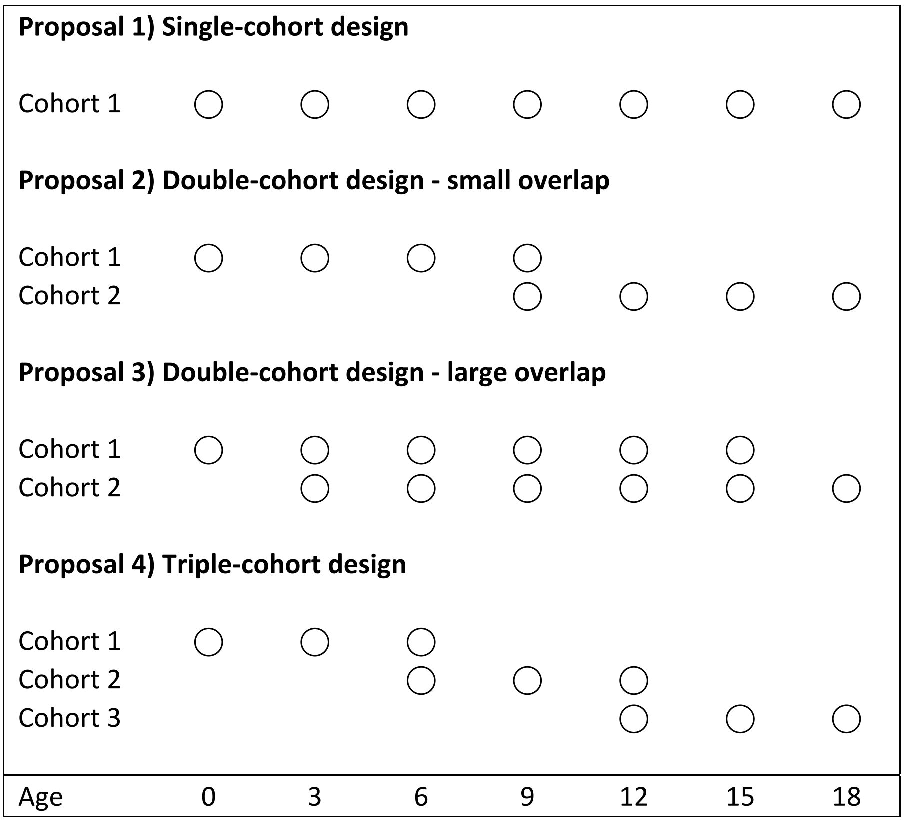
PAIRFAM used a design with three cohorts, with births spaced a decade apart (1971-3, 1981-3, 1991-3), and started interviewing in 2008/9 (Huinink et al. 2011). It aims to cover a much wider population than the cohort studies of children, and in this case the ALD is a way to gather cross-sectional information which is immediately useful as well as setting up the study.

8 Proposed designs

The project scope specifically outlines a specific focus on a new nationally representative birth cohort that spans from birth (or pregnancy) up to age 18+ years. With this in mind we review four possible designs and briefly discuss their pros and cons. The presented designs are suggested based on suggestions in the literature as well as common designs previously used. These are general guidelines and in practice a study design should be more specific to the study aims. The proposals are shown in Figure 2 with the time between measurements arbitrarily chosen to be every three years, over the suggested age range of 0-18 years old.

The first of the four proposed designs is a single-cohort design without any accelerated component, acting as a benchmark. The benefit of such a design for the UK is that it is consistent with previous national single-cohort studies from the 1958 National Child Development Study to the Millennium Cohort Study. Another single-cohort study, comparable to the previous ones will allow for cross-cohort studies. However, as discussed, the single-cohort studies do not provide generalisability to the national population as multi-cohort studies can.

Figure 2 Proposed cohort designs



The second and third proposed designs both have two cohorts. This was the recommended number of cohorts based on the study by Moerbeek (2011), albeit ignoring dropout. Hence if dropout is not expected to be high, then two cohorts are a desirable number. If two cohorts are considered, as in proposal 2 and 3, then the extent of overlap is a significant factor.

An overlap of one wave was suggested as a good compromise by Moerbeek (2011). This offers a balance to the trade-off since the more overlap the more power but also the more measurements. So in cases where power is less of an issue, but there are significant time and money constraints then a small overlap is preferable such as with proposal 2. A similar design was adopted for the GUI and the YOUth Study in Utrecht. For larger overlaps like in proposal 3, the power is stronger and the cohort effects will also be more distinguishable. This leads to more generalisable inference about the age effects. The study duration is longer, and more measurements are required for this proposal. Similar design to proposal 3 were adopted by the Australian LSAC, and GUS.

The fourth proposal is a triple-cohort design. This design would be advantageous if participant burden was high causing high dropout. This is because the more cohorts, the fewer waves are needed per subject. Furthermore, the power of the design is higher compared to two cohorts but not by substantial amounts (Moerbeek, 2011). For designs with three or more cohorts the logistics and management of the study become more demanding, as do power and sample size calculations. This design is somewhat similar to the Swiss COCON Study.

We make a point to again emphasise that a longitudinal study design should be aligned with the study aims, and these four proposals only offer general guidelines for general scenarios.

8.1 Practical considerations

There are many trade-offs in a design process, and for a general-purpose design it is always a compromise among the different objectives. Where recruitment is challenging, the single cohort design provides the most information for the least recruitment, but then sample units remain in the survey for the longest time to cover the age range, which increases the overall attrition and means that researchers have to wait longer for suitable data. The ALD reduces the impact of attrition and provides results faster, but requires recruitment of more sample units and potentially requires more complicated analytical models allowing for cohort effects.

Minimising the number of measurements leads to a cross-sectional design without repeated measures, with models derived from the observed age differences, but this is a risky design, because it relies on there being no cohort effects to be confounded with age effects. A longitudinal design allows age effects to be estimated more effectively, and this is why cohort studies are important. A random effects model such as (1) which allows these components to be separated needs to be estimable, and for ALDs with more cohorts and fewer observations per cohort there may be challenges over convergence (Galbraith *et al*. 2017). Equally, short sequences (in the extreme of two observations) are not suitable for estimating nonlinear age effects; to increase the possibility to identify nonlinear effects (which would be important in a multipurpose survey where it would be odd for all effects to be linear) longer sequences would be preferable. Moerbeek (2011)'s suggestion of two cohorts, based on the case of no attrition, nevertheless appears to be a reasonable starting point among these competing criteria.

The management of a data collection operation with more than one cohort also raises its own challenges. Each cohort potentially requires its own approach to data collection because it will contain respondents of different ages. This means that much more of the development work needs to be done initially for the survey to be released into the field. This may have benefits in consistency if the developments can go together, but this dies require sufficient specialist staff.

8.2 Sample size

Without a clear requirement for a cohort study and some decisions about aspects of the design like the number of cohorts in an ALD and the number of measurement occasions, it is quite difficult to make any recommendation on sample size. It is however interesting to consider the power curves in Galbraith et al. (2017, Fig. 4a, b), which suggest that around 80 sample units would give 90% power to detect a cohort effect in a ALD with two cohorts, and around 120 sample units for the same power in a with one cohort. These sizes are quite modest, but they apply equally to subpopulations of interest, so the recruitment would need to ensure these kinds of sizes in specific groups.

For estimation of other parameters and models the power and sample size requirements are more challenging; Galbraith et al. (2017) again provide some indication, showing the effect on the determinant of the fixed effects covariance matrix of variations in numbers of observations per sample unit in fixed size ALDs. The problem of finding the minimum sample size to achieve a fixed variance under ALDs is still open, and further research would be valuable on a range of different output variables.

9 Conclusion

Galbraith et al. (2017) provide a very useful exploration of the main axes of choice in a cohort design or ALD. Their analysis is based on the assumption that the linear model is correct, but their results are qualitatively similar under a quadratic model for the age effect, which suggests that they may have some robustness. For various practical reasons, including the practicality of managing the collection design, recruitment and field work for ALDs, a relatively small number of cohorts (two or three) seems to give the best trade-off. It opens up the possibilities to model cohort effects, and is well within the trade-offs between different aspects of the methodology for analysis.

More research is needed on determining sample sizes to meet different design constraints. The available design information is based on single variables, so there is also a need to balance the requirements across a wider range of variables and domains to produce a general purpose design which can meet multiple analytical needs.

10 References

Bell, R. (1953). Convergence: An Accelerated Longitudinal Approach. Child Development, 24(2), 145-152. http://dx.doi.org/10.2307/1126345

Bell, R. (1954). An Experimental Test of the Accelerated Longitudinal Approach. Child Development, 25(4), 281-286. http://dx.doi.org/10.2307/1126058

Buchmann, M. (2017). The Swiss Survey on Children and Youth (COCON). Presentation notes from the the Jacobs Center Workshop on Longitudinal Studies at the University of Zurich. https://www.ccwd.uzh.ch/assets/documents/research/knowledge-products/summaries/buchmann-summary.pdf

Buscot, M.-J., Wotherspoon, S. S., Magnussen, C. G., Juonala, M., Sabin, M. A., Burgner, D. P., Lehtimäki, T., Viikari, J. S. A., Hutri-Kähönen, N., Raitakari, O. T. & Thomson, R. J. (2017). Bayesian hierarchical piecewise regression models: a tool to detect trajectory divergence between groups in long-term observational studies. BMC Medical Research Methodology, 17(1), 86.

Curran, P. J. (2003) Have Multilevel Models Been Structural Equation Models All Along? Multivariate Behavioral Research, 38:4, 529-569, http://dx.doi.org/10.1207/s15327906mbr3804_5

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve Frequently Asked Questions About Growth Curve Modeling. Journal of Cognition and Development : official journal of the Cognitive Development Society, 11(2), 121–136. http://dx.doi.org/10.1080/15248371003699969

Duncan, S. C., Duncan, T. E., & Hops, H. (1996). Analysis of longitudinal data within accelerated longitudinal designs. Psychological Methods, 1(3), 236-248. http://dx.doi.org/10.1037/1082-989X.1.3.236

Galbraith, S., Bowden, J., & Mander, A. (2017). Accelerated longitudinal designs: An overview of modelling, power, costs and handling missing data. Statistical methods in medical research, 26(1), 374–398. http://dx.doi.org/10.1177/0962280214547150

Gray, M., & Sanson, A. (2005). Growing up in Australia: The longitudinal study of Australian children. Family Matters, (72), 4. https://aifs.gov.au/sites/default/files/mg(2).pdf

Harring, J.R., McNeish, D.M., & Xiaoshu (2016). On the Adequacy of SEM Model Fit Criteria to Detect Cohort Effects in Accelerated Longitudinal Designs.

Huinink, J., Brüderl, J., Nauck, B., Walper, S., Castiglioni, L., & Feldhaus, M. (2011). Panel analysis of intimate relationships and family dynamics (pairfam): Conceptual framework and design. Journal of Family Research, 23(1) 77-100.

Michaud, S. (2001). The National Longitudinal Survey of Children and Youth-Overview and Changes After Three Cycles. Canadian Studies in Population, 28(2), 391-405. https://doi.org/10.25336/P6QK58

Miyazaki, Y., & Raudenbush, S. W. (2000). Tests for linkage of multiple cohorts in an accelerated longitudinal design. Psychological methods, 5(1), 44-63.

Moerbeek, M. (2008). Powerful and Cost-Efficient Designs for Longitudinal Intervention Studies With Two Treatment Groups. Journal of Educational and Behavioral Statistics, 33(1), 41–61. https://doi.org/10.3102/1076998607302630

Moerbeek, M. (2011). The effects of the number of cohorts, degree of overlap among cohorts, and frequency of observation on power in accelerated longitudinal designs. Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 7(1), 11-24. http://dx.doi.org/10.1027/1614-2241/a000019

Onland-Moret, N. C. (2014). Substantiation measurement at fixed- or variable ages. Unpublished summary of discussions.

Raudenbush, S. W., & Liu, X.-F. (2001). Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. Psychological Methods, 6(4), 387-401. http://dx.doi.org/10.1037/1082-989X.6.4.387

Schaie, K. W., & Strother, C. R. (1968). A cross-sequential study of age changes in cognitive behavior. Psychological Bulletin, 70(6, Pt.1), 671-680. http://dx.doi.org/10.1037/h0026811

Thurber, K. A., Banks, E., Banwell, C., & LSIC Team (2015). Cohort Profile: Footprints in Time, the Australian Longitudinal Study of Indigenous Children. International journal of epidemiology, 44(3), 789–800. http://dx.doi.org/10.1093/ije/dyu122

Wu, W., West, S. G., & Taylor, A. B. (2009). Evaluating model fit for growth curve models: Integration of fit indices from SEM and MLM frameworks. Psychological Methods, 14(3), 183-201.