Kingston A, Jagger C.
[Review of methodologies of cohort studies of older people](https://doi.org/10.1093/ageing/afx183).
Age and Ageing 2017, 47(2), 215-219

DOI link to article:

Date deposited:

28/02/2018

Embargo release date:

01 December 2018

Review of methodologies of cohort studies of older people

Andrew Kingston[1,2], Carol Jagger[1,2]

[1]Instiute of Health and Society, Newcastle University, United Kingdom

[2]Newcastle University Institute for Ageing

Abstract

This paper describes a cohort study in terms of its design, the research questions answered by cohort studies, common analytic techniques, and the strengths and limitations of this type of study. We also describe the main cohort studies of older populations, many of which are available for secondary data analysis.

Introduction

Compiling evidence for cause and effect is the basis of much ageing research. The best evidence comes from randomised controlled trials, but this design is not always possible in practice, particularly where the 'cause' is potentially harmful, for instance smoking or obesity. The next best evidence is then from observational cohort studies. In this article we describe the design of cohort studies including the type of research question they can answer, common analytical techniques for cohort studies, and the strengths and limitations of this type of design. Finally, we provide a brief description of the major cohort studies of older populations, many of which are available to researchers outside of the study team for secondary data analysis. Comparative analysis of cohort studies from different populations can considerably strengthen evidence of cause and effect, or, in some cases, can help tease out mediating factors available in one, but not the other cohort.

Design of cohort studies

A cohort study is simply a set of individuals who share some characteristic and who are followed up over time. In terms of cohort studies of ageing this is generally birth year (or years), but it might be another shared characteristic such as occupation or geography. The majority of ageing research is fundamentally trying to determine whether there is a relationship between a cause or exposure, and an effect or outcome. The 'exposure' might be naturally occurring, a behaviour, or a characteristic. Examples of research questions answered by cohort studies include 'Does a healthy diet (behavioural exposure) reduce the risk of dementia (outcome)?' or 'Does childhood socioeconomic status (characteristic – exposure) predict midlife psychological distress (outcome)?'. In some cases, the exposure might be a service or intervention that cannot be evaluated by a RCT, often because the service is already in place and withdrawal for some would be unethical or impossible.

In all cases the exposure must happen before the outcome to infer cause and effect. Hence the natural starting point is with the exposure – either observe outcome(s) that result from a given exposure (observational cohort study) or control the exposure and record the outcome(s) (RCT). Thus, if we were researching the effect of vitamin supplementation on a disease, we can contrast the design for a RCT and a cohort study:

- In a RCT the investigator will decide which participants receive vitamin supplements and which do not, and compare outcomes for the two groups;
- In a cohort study, people who happen to be taking vitamin supplements are compared to those who do not. i.e there is no active intervention by the investigator.

The two designs can also be contrasted in terms of the timing of the selection of participants and the occurrence of the exposure (Figure 1).

Observing the occurrence of the outcome may not necessarily involve collecting data directly from study members, for example mortality or cancer incidence, as well as certain other diseases, as these may be ascertained from linkage with administrative databases or registers. However, since cohort studies have the capacity to investigate multiple outcomes, many cohort studies collect data directly from study members at regular follow-ups. For ageing cohorts, follow-up intervals may reduce as the cohorts reach very old age in order not to 'miss' outcomes through death.

Strengths and limitations of cohort studies

The major strength of cohort studies is that they follow the natural course of events so there is no confusion about whether the exposure preceded the outcome or vice versa. Multiple outcomes may be investigated, and this is the case for the major cohort studies of ageing. Unlike case-control studies, cohort studies avoid bias in exposure measurement as the outcome is unknown when information on the exposure is ascertained.

Limitations of cohort studies include the large number of participants and the generally long follow-up period required, and therefore cohort studies are costly. Many existing cohort studies of ageing sample only from community-dwelling older adults and are therefore not representative of the total population. This is obviously a limitation for estimating

prevalence and incidence, but it may also be an issue when investigating causal mechanisms [1]. There is also the problem of loss to follow-up, often through death of study members before outcomes occur, which means the cohort becomes less representative over time. When cohort studies require long follow-up periods there is the possibility that diagnostic criteria, methods and optimal exposure measurement may change over time. Hence there is sometimes the view that the answer to the research question has arrived too late.

Common analytical techniques for cohort studies

Most research from cohort studies attempts to determine the association between an exposure (or risk factor) and an outcome, and to quantify the magnitude of the effect of the risk factor on the outcome. The simplest method for this is to calculate the risk or rate ratio for the outcome in the exposed cohort compared to the risk or rate in the unexposed cohort. The denominator for the risk is the whole exposed (or not exposed) cohort whilst for the rate it is the person years of observation, which considers any losses to follow-up.

In a RCT, the process of randomisation generally (though not always) results in the exposed and non-exposed groups being balanced on all other factors that might be associated with both the exposure and the outcome. However, in cohort studies, the analysis must try to account for this by controlling, or adjusting, for such confounding factors that might obscure the effect of the exposure and outcome. Potential confounding variables are generally detected through causal diagrams and/or after careful thought of the causal pathways that govern the outcome of interest, including previous research. To account for confounding the simple comparison of the risk in the exposed and not exposed is extended to a logistic regression model where the dependent variable is whether or not the outcome has occurred and exposure is the independent variable. Other potential confounding factors are then added in to see whether these affect the relationship between exposure and outcome. However, care must be taken that other factors are not mediators, or part of the causal pathway between the exposure and the outcomes.

When the outcome is continuous (e.g. gait speed) and measured at baseline and follow-ups then the analysis must take account of the dependence between the multiple observations on the same individual. For balanced data (no missing observation and equally spaced

follow-ups, the simplest analysis is a repeated measures analysis of variance. However, such repeated measures can be fitted much more flexibly in the multi-level modelling framework which can accommodate outcomes that are binary, ordinal or even time-to-event (see below) as well as continuous. Training material for multi-level models is available from the Bristol Centre for Multilevel Modelling (http://www.bristol.ac.uk/cmm/) but these models can also be fitted in standard statistical software packages such as the Statistical Package for the Social sciences (SPSS).

Some cohort studies are interested in how long it takes for an outcome/event to occur and this is often referred to as time-to-event data, survival data or failure data. In many cases the time to event is death. To analyse these types of data, a routinely observed test to detect differences in survival rates between groups is the log-rank test (not accounting for confounding). Survival rates between two groups can be compared through graphical plots of the Kaplan Meir estimate of the survivor function and/or Nelson-Aalen estimate of cumulative hazard. If there is a need to assess the effects of an exposure after accounting for confounding, the most commonly used modelling technique is the Cox Proportional Hazards model. Other techniques can be used but have additional assumptions.

These techniques outline of the main statistical methods that can be found in the cohort study literature but are not exhaustive. Other more specialist techniques are available, for example multi-state modelling techniques when the outcome/event is not binary (dead/alive) but a series of health states (for example independent/disabled/dead) observed on multiple occasions over time. Most statistical textbooks will cover the basic methods but complex modelling of longitudinal data would benefit from a more in-depth approach [2].

Major cohort studies of ageing

It is impossible here to review all the cohort studies which include older populations, or which were developed to answer research questions on ageing. Here we focus on studies which have been developed to answer general research questions on ageing, rather than those focused on specific chronic conditions such as osteoporosis, diabetes or

cardiovascular disease. General ageing cohorts can be divided into two groups: those covering an age range, and those recruiting participants from a single birth year. An important subgroup of the former are the growing number of studies based on and harmonised with the long running US Health and Retirement Study [3], including the English Longitudinal Study of Ageing (ELSA) [4], the Irish Longitudinal Study on Ageing (TILDA) [5], the Survey of Health and Retirement in Europe (SHARE) [6], and the China Health and Retirement Longitudinal Study (CHARLS) [7]. Data from HRS and many of its international sister studies are publicly available and the full list of studies, with tools and resources from cross harmonisation, are available on the USC Gateway to Global Aging (www.https://g2aging.org/).

Other long running cohort studies of ageing covering more than a single birth cohort have generally been regional rather than national; these include the Longitudinal Ageing Study Amsterdam (LASA) [8], the Jerusalem Longitudinal Study [9], the Cambridge City over-75s Cohort (CC75C) [10], the Cognitive Function and Ageing Studies [11, 12], and a set of Australian studies which have been brought together and harmonised in the Dynamic Analyses to Optimize Ageing (DYNOPTA) project [13]. Rather than geographically based, others are occupationally focussed such as the Whitehall Study [14], the GAZEL cohort [15], and the Helsinki Health Study [16].

Cohort studies recruiting participants from a single birth year can also be subdivided into those who recruit from birth or childhood, thereby collecting early life information without recall bias, and those who recruit in old age. The UK is fortunate in having a rich history of cohorts followed regularly from birth and the first of these, the 1946 National Birth Cohort [17], has now reached older age (65+ years) and the second, the 1958 British Birth Cohort, will reach 60 in 2018 [18]. Two other British cohorts had data from birth records but were not retraced until later in life: the Newcastle Thousand Families Study [19], and the Hertfordshire Cohort Study [20], whereas studies such as the Lothian Birth cohorts of 1921 and 1936 [21], the Boyd Orr cohort [22], and the Aberdeen Children of the 1950s cohort [23], retraced in adulthood those who took part in surveys or tests as children, thereby providing early life data.

The very old, those aged 85 years and over, are the fastest growing section of most populations in the developed world. Cohort studies of ageing which have a lower age limit of 50 or 65 years often contain relatively few very old people, unless this age group is over sampled (as in CFAS [11]). To overcome this there has been a set of studies with similar design and based on single year birth cohorts but recruited at very old age: the Leiden 85+ cohort [24], the Newcastle 85+ study [25], and the Life and Living in Advanced Age Cohort Study in New Zealand (Lilacs NZ) [26]. Other cohort studies (not necessarily single birth years) of the very old include the Vitality 90+ study from Finland [27], the Swedish Panel Study of Living Conditions of the Oldest Old (SWEOLD) [28], and the various centenarian cohorts [29].

The list of cohort studies of ageing cited here is not exhaustive. Outcomes usually cover functional and cognitive health, but many would also encompass social functioning. Increasingly objective as well as subjective health and functioning measures are collected, including biological measures. Cohorts from the Nordic countries especially, and increasingly the UK, link administrative data or data on health and social care usage and, for younger members, employment. Given the expense of set up and maintenance of cohort studies, there are a growing number of initiatives to meta-analyse cohort studies through post harmonisation, such as IALSA (Integrative Analysis of Longitudinal Studies of Aging and Dementia, available at https://staging.maelstrom-research.org/mica/network/ialsa/) and the UK based CLOSER (Cohort & Longitudinal Studies Enhancement Resources, available at https://www.closer.ac.uk/).

Of the studies outlined here some features are worthy of mention:
i)   Population-based samples do not always include those living in assisted care facilities (residential or nursing homes); in studies of the very old this can be problematic and can give a distorted picture of the health of this age group.
ii)   Although studies followed from birth have data from across the life course, they require very long-term funding and may be relatively underpowered by advanced age.

iii)   Methods of measuring health in younger old populations may not be appropriate for very old populations leading to biased results; in particular measures may have ceiling or floor effects at younger or very old ages.

iv)   Though clinic based assessments may provide more accurate assessment, they have the potential to introduce bias since participants with significant mobility or cognitive impairments may not take part [30].

v)   Maintaining active contact with participants between assessment waves is crucial to reduce withdrawal from the study, as well as for ascertainment of address change or death which can reduce unnecessary or inappropriate contact.

Conclusions

If well conducted with minimal attrition, cohort studies can provide strong evidence of causal links between an outcome and an earlier exposure. Many cohort studies of older populations already exist but they are expensive to set up and maintain, and therefore new longitudinal cohorts are being focussed in countries where none exist. The HRS set of studies, with similar designs and measures, maximise the potential to validate findings across countries with different cultures, societal norms and health delivery, thereby strengthening evidence. Other post harmonisation efforts, especially through meta-analysis, are important and should be employed more often to provide stronger evidence for policy in the shortest time.
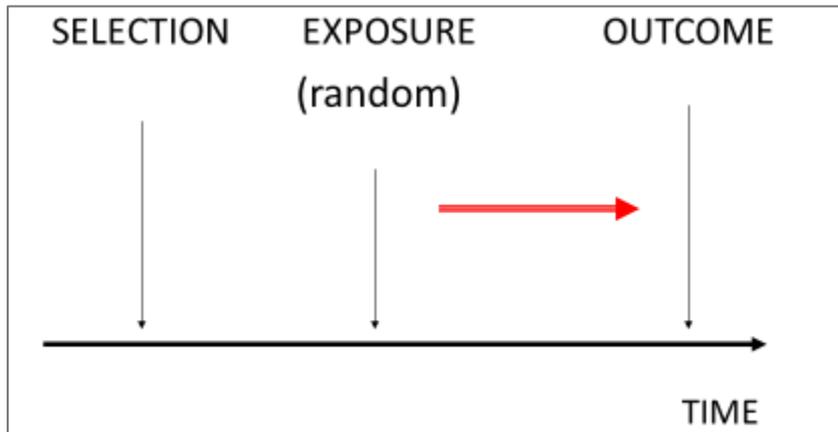
REFERENCES

1.   Kelfve, S., *Underestimated Health Inequalities Among Older People-A Consequence of Excluding the Most Disabled and Disadvantaged.* J Gerontol B Psychol Sci Soc Sci, 2017.
2.   Singer, J.D. and J.B. Willett, *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. 2003, Oxford: Oxford University Press.
3.   Sonnega, A., et al., *Cohort Profile: the Health and Retirement Study (HRS).* International Journal of Epidemiology, 2014. **43**(2): p. 576-585.
4.   Steptoe, A., et al., *Cohort Profile: The English Longitudinal Study of Ageing.* International Journal of Epidemiology, 2013. **42**(6): p. 1640-1648.
5.   Kearney, P.M., et al., *Cohort Profile: The Irish Longitudinal Study on Ageing.* International Journal of Epidemiology, 2011. **40**(4): p. 877-884.

6.   Börsch-Supan, A., et al., *Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE).* International Journal of Epidemiology, 2013.

7.   Zhao, Y.H., et al., *Cohort Profile: The China Health and Retirement Longitudinal Study (CHARLS).* International Journal of Epidemiology, 2014. **43**(1): p. 61-68.

8.   Huisman, M., et al., *Cohort Profile: The Longitudinal Aging Study Amsterdam.* International Journal of Epidemiology, 2011. **40**(4): p. 868-876.

9.   Jacobs, J.M., et al., *Cohort Profile: the Jerusalem longitudinal cohort study.* International Journal of Epidemiology, 2009. **38**(6): p. 1464-1469.

10.   Fleming, J., et al., *Cohort profile: The Cambridge City over-75s Cohort (CC75C).* International Journal of Epidemiology, 2007. **36**(1): p. 40-46.

11.   Brayne, C., et al., *Cohort profile: the Medical Research Council Cognitive Function and Ageing Study (CFAS).* International Journal of Epidemiology, 2006. **35**(5): p. 1140-5.

12.   Matthews, F.E., et al., *A two-decade comparison of prevalence of dementia in individuals aged 65 years and older from three geographical areas of England: results of the Cognitive Function and Ageing Study I and II.* Lancet, 2013. **382**(9902): p. 1405-1412.

13.   Anstey, K.J., et al., *Cohort profile: The Dynamic Analyses to Optimize Ageing (DYNOPTA) project.* International Journal of Epidemiology, 2010. **39**(1): p. 44-51.

14.   Marmot, M. and E. Brunner, *Cohort profile: The Whitehall II study.* International Journal of Epidemiology, 2005. **34**(2): p. 251-256.

15.   Goldberg, M., A. Leclerc, and M. Zins, *Cohort Profile Update: The GAZEL Cohort Study.* International Journal of Epidemiology, 2015. **44**(1): p. 77-U102.

16.   Lahelma, E., et al., *Cohort Profile: The Helsinki Health Study.* International Journal of Epidemiology, 2013. **42**(3): p. 722-730.

17.   Wadsworth, M., et al., *Cohort profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development).* International Journal of Epidemiology, 2006. **35**(1): p. 49-54.

18.   Power, C. and J. Elliott, *Cohort profile: 1958 British Birth Cohort (National Child Development Study).* International Journal of Epidemiology, 2006. **35**(1): p. 34-41.

19.   Pearce, M.S., et al., *Cohort Profile: The Newcastle Thousand Families 1947 Birth Cohort.* International Journal of Epidemiology, 2009. **38**(4): p. 932-937.

20.   Syddall, H.E., et al., *Cohort profile: The Hertfordshire Cohort Study.* International Journal of Epidemiology, 2005. **34**(6): p. 1234-1242.

21.   Deary, I.J., et al., *Cohort Profile: The Lothian Birth Cohorts of 1921 and 1936.* International Journal of Epidemiology, 2012. **41**(6): p. 1576-1584.

22.   Martin, R.M., et al., *Cohort Profile: The Boyd Orr cohort - an historical cohort study based on the 65 year follow-up of the Carnegie Survey of Diet and Health (1937-39).* International Journal of Epidemiology, 2005. **34**(4): p. 742-749.

23.   Leon, D.A., et al., *Cohort profile: The Aberdeen children of the 1950s study.* International Journal of Epidemiology, 2006. **35**(3): p. 549-552.

24.   der Wiel, A.B.-v., et al., *A high response is not essential to prevent selection bias.* Journal of Clinical Epidemiology. **55**(11): p. 1119-1125.

25.   Collerton, J., et al., *The Newcastle 85+ study: biological, clinical and psychosocial factors associated with healthy ageing: study protocol.* BMC Geriatrics, 2007: p. 7-14.

26.   Kerse, N., et al., *Cohort Profile: Te Puawaitanga o Nga Tapuwae Kia Ora Tonu, Life and Living in Advanced Age: a Cohort Study in New Zealand (LiLACS NZ).* International Journal of Epidemiology, 2015. **44**(6): p. 1823-1832.

27.   Jylhä, M. and A. Hervonen, *Functional status and need of help among people aged 90 or over: a mailed survey with a total home-dwelling population.* Scandinavian Journal of Public Health, 1999. **27**(2): p. 106-111.

28.   Lennartsson, C., et al., *Data resource profile: The Swedish Panel Study of Living Conditions of the Oldest Old (SWEOLD).* Int J Epidemiol, 2014. **43**(3): p. 731-8.

29.   Willcox, D.C., B.J. Willcox, and L.W. Poon, *Centenarian Studies: Important Contributors to Our Understanding of the Aging Process and Longevity.* Current Gerontology and Geriatrics Research, 2010. **2010**: p. 6.

30.   van Bemmel, T., et al., *Impact of valvular heart disease on activities of daily living of nonagenarians: the Leiden 85-plus study a population based study.* BMC geriatrics, 2010. **10**(1): p. 1.

Figure 1: Design of randomised controlled trial (RCT) and observational cohort study

Randomised Controlled Trial (RCT)

Cohort Study