Analytical approaches in cross-cohort research

Graciela Muniz-Terrera
Programme Leader Track
MRC Unit for Lifelong Health and Ageing
University College London
September
2015

Cross-cohort work approaches

- Data pooling
  - Aggregate data meta analysis (AG)
  - Individual patient meta analysis (IPD)

IPD increases power to detect treatment effects (particularly useful in randomized clinical trials); technically challenging

- Integrative/ independent data analysis (IDA)
  - Coordinated analytical approach (Piccinin et al 2013)

Useful for evaluation of replicability & of patterns of results

Cross-cohort work approaches

- Data pooling and IDA require different levels of data harmonisation
- Harmonisation: qualitative or statistical
  - Qualitative: ex creating study specific study cut points for variables like age to convert data into common format
  - Statistical: uses specialised methods to derive common format data
  - Unfortunately, neither is common practice in systematic reviews
- Harmonization: prospective or retrospective
  - Prospective: data collected following a previously agreed protocol (ex: SHARE)
  - Retrospective: done after data has been collected, not always possible or complete (can be partial)

Data pooling

Data pooling

- Data pooling, either for AD or IDA, requires data harmonisation to obtain interpretable results
- When pooling cognitive data, researchers are faced with multiple challenges, as outcomes are rarely binary
- Quite often, cognitive tests consist of a questionnaire such that correct answers to individual questions are scored and the final test score is the sum of the individual scores
- For example, to evaluate global cognitive function, many studies use the Mini Mental State Exam

Mini Mental State Examination

Mini-Mental State Examination (MMSE)

Patient's Name: ______________________________ Date: ______________

Instructions: Score one point for each correct response within each question or activity.

Maximum Score	Patient's Score	Questions
5		"What is the year? Season? Date? Day? Month?"
5		"Where are we now? State? County? Town/city? Hospital? Floor?"
3		The examiner names three unrelated objects clearly and slowly, then the instructor asks the patient to name all three of them. The patient's response is used for scoring. The examiner repeats them until patient learns all of them, if possible.
5		"I would like you to count backward from 100 by sevens." (93, 86, 79, 72, 65, …) Alternative: "Spell WORLD backwards." (D-L-R-O-W)
3		"Earlier I told you the names of three things. Can you tell me what those were?"
2		Show the patient two simple objects, such as a wristwatch and a pencil, and ask the patient to name them.
1		"Repeat the phrase: 'No ifs, ands, or buts.'"
3		"Take the paper in your right hand, fold it in half, and put it on the floor." (The examiner gives the patient a piece of blank paper.)
1		"Please read this and do what it says." (Written instruction is "Close your eyes.")
1		"Make up and write a sentence about anything." (This sentence must contain a noun and a verb.)
1		"Please copy this picture." (The examiner gives the patient a blank piece of paper and asks him/her to draw the symbol below. All 10 angles must be present and two must intersect.) 
30		TOTAL

Data pooling

- Questionnaire data in different studies may have missing items

(MMSE may not include some questions, some say too loaded on memory)

- Language differences may be considerable & translation of tests is not direct.
- Difficulty of the items may vary between countries
- Even within languages variations should be expected (ex: arete, caravana, pendiente; medias, calcetas, escarpines, calcetines)
- These differences are usually ignored when pooling data

Methods used in harmonization of cognitive variables

- Ignoring item harmonization, various methods have been proposed.
- Each method makes different assumptions, we will discuss them when necessary
- Three commonly used methods of harmonization are:
  - T-scores
  - C-scores
  - Calculation of latent variables

T-scores

- Converts raw cognitive measures to demographically corrected standardized scores
  - Normalize each score (ex: wrt mean 10, st dev 3)
  - Regress each t-scores on age, sex, education (wrt. specific group by centering)
  - Calculate residuals=(actual scaled score-predicted scaled score)
  - Convert residuals to t-scores
  - These t-scores are interpreted as how an individual's score on each cognitive measure compares to the average score of participants of same sex and age & education.

C-scores

- Calculation of study specific scores standardized relative to a consistent group across datasets
  - sociodemographic information used to identify the centering group
  - Ex: C-score=(raw score-mean females aged 70-74, with 8 yrs education / st. dev of females aged 70-74, with 8 yrs. education)
- C-scores & t-scores do not take into account the differences between the measurements properties of the scales

Latent variable approach

- The LVM assumes that the overall test score of a participant is influenced by a univariate continuous variable unique to that participant
- The overall test score is viewed as counts representing a correct number of scored test item
- Conditional on the latent variable, the overall test score follows a binomial distribution such that:

$$Y_{ij} = y_{ij} \mid Z_i = z \sim Bin(N_j, p_{ij}(z)); Z_i \sim N(0, \tau_i^2)$$

Y_{ij}= number of correctly scoring test items for test for person *i*

Probability of correctly scoring an item is modelled using a logistic regression

$$\text{logit}(p_{ij}(z)) = \beta_{0j} + \sum_{k=1}^{K} \beta_k x_{ik} + z$$

where the intercept is related to the difficulty of the test

Latent variable approach

- The st. deviation τ_i^2 can also be related to covariates and indicates whether the tests can discriminate between groups (whether performance differences between men are larger than between women); such that

$$\log(\tau) = \eta_0 + \sum_{k=1}^{K} \eta_k x_{ik}$$

- An essential assumption of this method is that the mean and variance of the latent variable for each item would be the same across studies (factorial invariance)
- This method is technically complex and requires the use of specialised software

Comparison of the 3 approaches

Griffin et al. compared the 3 methods and concluded that:

- T-scores least desirable compared to C-scores or latent variable (LV) method.
- T-scores assume that the sample characteristics are identical across studies. C-scores makes similar assumption but for a subgroup.
- Within study standarisation (T-scores) is not desirable as it ignores differences in scale and sample distributions
- LV most desirable & only method that allows for testing measurement invariance
- Higher complexity when longitudinal data is harmonized

Assumptions of different methods

Table 5. Assumptions for the different classes of statistical harmonization methods

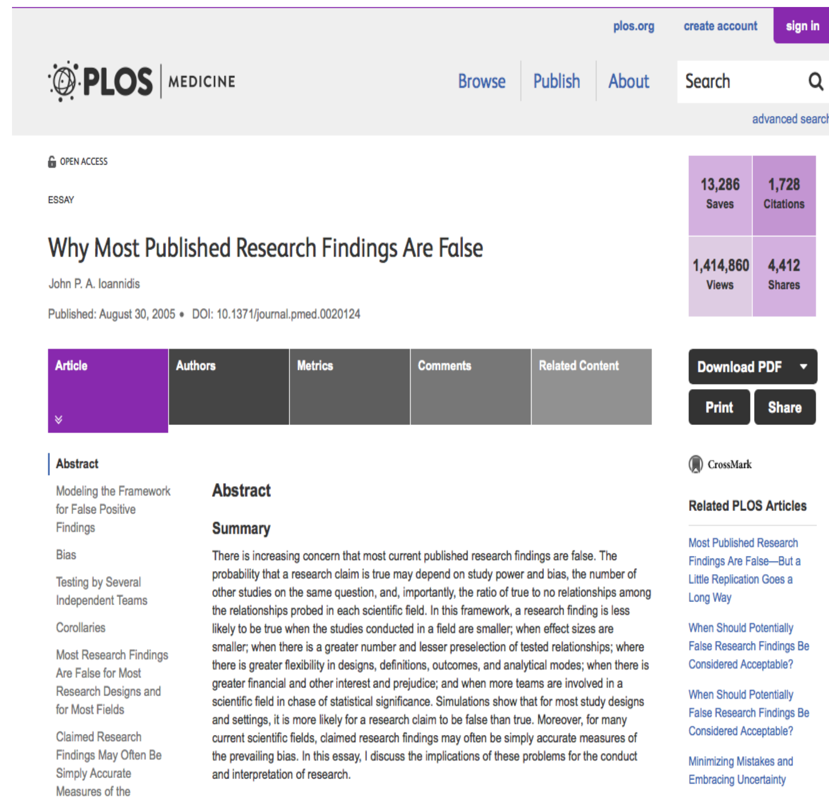
Method	Assumptions	How Can It Be Applied
Standardization Methods 6 studies used this class of methods, e.g., Minicuci, N. 2003[74]	• Scales have an underlying normal distribution • The scales have a similar distribution (i.e., being in the 5^{th} percentile of one scale is equivalent to being in the 5^{th} percentile of another)	Can be applied in most situations with continuous variables and does not require specialized software Does not require common items across studies Need to transform back to a chosen scale(s) for interpretation
Item Response Theory Latent Variable Model 15 studies used this class of methods, e.g., Van Buuren, S. 2005;[76] Bauer, DJ. 2009;[70] McArdle, J. 2009[73]	• Underlying constructs are unidimensional • Some items must be common across datasets or at least can be "chained" together • The items are equally discriminating (only for IP and Rasch models) • Factorial invariance If repeated measures: • Item difficulty is invariant with respect to time or age • Item discrimination does not change across time or age	Can be applied to continuous, binary and ordinal data but requires some specialized software Can accommodate different scale types among items However can be extended to include longitudinal data as per McArdle, et al. by integrating IRT and latent curve modeling using a joint model likelihood approach
Missing data by design with multiple imputation 3 studies used this class of methods, e.g., Burns, RA. 2011[71]	• Missingness is assumed to be at random (i.e., MAR) • Some items must be common across datasets or at least can be "chained" together	Can be applied to continuous, binary and ordinal data but requires some specialized software and multiple datasets Can accommodate different scale types among items Can be used if scales are not unidimensional

IRT = item response theory; MAR = missing at random

Integrative data analysis (IDA)

Coordinated analytical approach

- Often, researchers are interested in replicability /reproducibility of results.



Example

A long standing question in ageing research is whether education is associated with cognitive decline.

Lenehan et al (2014), in a recent review published:

Table 1 Studies examining the association between education and age-related cognitive decline

Authors and date	*n*	Sample and sampling method	Age range at baseline (years)	Education categorical or continuous	Mean education ± SD (years)	Study design and analysis	Cognitive functions	Findings
Cullum *et al.* (2000)[34]	135	Subsample of the Cambridge City Over-75 Cohort; a population-based sample drawn from general practice lists	75–85+	Categorical: <15 years (64%) >15 years (36%)	NA	Two assessments over 4 years; logistic regression	The Cambridge Cognitive Examination subscales: memory, attention/calculation (combined), perception, orientation, praxis, abstract-thought and language	Less education is associated with decline in memory subscale only. Declines occurred in all other functions but were not associated with education.
Christensen *et al.* (2001)[25]	887	Canberra Longitudinal Study; probability sample of persons aged over 70 years drawn from electoral roll, community dwelling	70–93	Continuous & <10 (*n* = 68) 10–12 (*n* = 127) >12 (*n* = 99)	NA	Three assessments over 8 years: (i) latent growth curve modelling; (ii) ANOVA; and (iii) regression analyses on survivors for whom complete data available (*n* = 294)	Crystallised intelligence (vocabulary, similarities, and NART); memory (word recognition, recall of three items, address recall); speed (SLMT); general cognitive function (MMSE)	Education was significantly related to level of CIQ, memory, and speed. Education level was not associated with differences in rates of decline on any cognitive measure. Education was associated with better performance in CIQ, speed, and MMSE, but not memory. Decline was evident across 8-year period for speed, memory, and MMSE, but not for CIQ; there were no differences in rate of decline as a function of level of education for any function.
Bosma *et al.* (2003)[35]	708	Maastricht Aging Study longitudinal data; convenience sample drawn from a registration network of general practices	50–80	Continuous and three categories ranging from primary education to university education	NA	Two assessments over 3 years; ordinary least squares regression	Processing speed (modified Stroop-Colour-Word Test); verbal memory (Verbal Learning Test); general cognitive function (MMSE)	Low educational level was associated with faster decline in speed, memory, and general cognitive function when compared to a high educational level. The associations lose statistical significance when controlling for mental workload and intellectual abilities.
Seeman *et al.* (2005)[33]	895	MacArthur Successful Aging Study data; population-based sample from which a subsample was drawn on the basis of age and physical and mental health	70–79	0–8 years (reference group) (29.1%) 9–11 years (25.5%) 12 years (24.0%) 13 + years (21.5%)	Overall 10.64 ± 3.43	Three assessments over 7 years; mixed models	Memory (sum of delayed incidental recall and delayed spatial recognition); abstraction (four items of similarities); language (modified BNT); spatial ability (figures); global cognition (sum of scores on five tests listed above)	Higher education was associated with better performance on all five cognitive measures. There were no significant differences in rates of decline as a function of education level across any function. For those with 13+ and 9–11 years education, the APOE ε4 allele was associated with faster decline in global cognition over time, similar to a trend observed for those with 12 years education ($P > 0.05$).
Alley *et al.* (2007)[24]	6651	Asset and Health Dynamics of the Oldest Old data; nationally representative sample of older Americans living in the community	70+	Continuous	11.1 ± 3.5	Four assessments over 7 years; growth curve modelling	Verbal memory (delayed and immediate recall); working memory (serial sevens); general mental status (Telephone Interview for Cognitive Status)	Higher education was related to better performance on all three cognitive tests. Higher education was associated with slower decline in general mental status and faster decline in verbal memory; it was unrelated to the rate of decline in working memory.
Van Dijk *et al.* (2008)[29]	872	MacArthur Successful Aging Study data; subsample drawn on the basis of age and cases with no missing data; convenience sample randomly drawn from general practice registers	49–81	Categorical: Low (primary and lower vocational, ≤10 years) High (secondary education or university)	Low: 8.3 ± 1.6 High: 11.3 ± 2.9 Total sample: 9.9 ± 2.8	Three assessments over 6 years; linear mixed modelling	Verbal learning (the Verbal Learning Test); long-term memory (delayed recall modified RAVLT); attention switching (modification of Trail Making); semantic fluency (verbal fluency test); phonemic fluency (verbal fluency test); interference control (Stroop-Colour-Word Test); mental speed (Letter Digit Substitution Test); general cognitive status (MMSE)	Higher education was related to better performance on all cognitive tests. Rate of decline did not differ depending on educational level on any of the cognitive tests.

Proust-Lima *et al.* (2008)[32]	1800	Personnes Agées QUID data, subsample without dementia; convenience sample randomly selected from electoral roll	65+	Categorical low (no primary school diploma) – ≤6 years of education (*n* = 453) High (primary school diploma) – ≥6 years education (*n* = 1347)	NA	Eight assessments over 15 years; non-linear latent process models	Global cognitive performance (MMSE); verbal fluency (Isaacs Set Test); verbal memory (recognition form of the Benton Visual Retention Test); psychomotor speed (Digit Symbol Substitution Test); latent cognitive factor (the common factor of the four psychometric tests)	Linear mixed models showed that subjects with higher education performed better on visual memory and psychomotor speed tasks, but there were no significant differences between education groups on MMSE or verbal fluency score. Subjects with higher education declined at a faster rate for measures of global cognitive function and psychomotor speed. Non-linear models revealed that higher education was associated with faster decline in verbal fluency and psychomotor speed. There were no significant differences in rate of decline between performance on the MMSE or visual memory task.
Tucker-Drob *et al.* (2009)[26]	690	Advanced Cognitive Training for Independent and Vital Elderly; subsample no-contact control group; convenience sample drawn from various registers and settings	65–94	Continuous	13.4 ± 2.7 (range: 6–20)	Five assessments over 5 years; latent growth curve modelling techniques	Reasoning (letter series, word series and letter sets); processing speed (three tasks from the field-of-view measure); vocabulary (a test from the Kit of Factor Referenced Cognitive Tests); composite test scores representing reasoning and speed were also computed.	Education was related to cognitive performance but not associated with rates of cognitive decline over time, both before and after baseline education was controlled for.
Der *et al.* (2010)[28]	398	Healthy Old People in Edinburgh study; subsample based on completion of cognitive tests; convenience sample identified through age registers of general practices	70+	Continuous	10.9 ± 2.6 at baseline	Three assessments over 9 years; linear mixed effects modelling	Fluid intelligence/non-verbal inductive reasoning (Raven's Standard Progressive Matrices) and verbal declarative memory (logical memory).	Participants with higher education had a higher mean score on both cognitive outcomes at baseline. There were no interaction effects between age and education, suggesting there are no differences between rates of cognitive decline between those with lower or higher education levels.
Zahodne *et al.* (2011)[27]	1014	Victoria Longitudinal Study; two subsamples based on follow-up period; convenience sample consisted of community-dwelling volunteers	54–95	Continuous and categorical ≤13 years or ≥14 years education	Sample 1: 13.4 Sample 2: 14.7 Entire sample: 14.1 ± 3.1 (range: 6–20)	Up to five assessments over 12 years; unconditional and conditional growth models	Verbal processing speed (lexical decision and sentence verification); working memory (sentence construction and two span tests); verbal fluency (three tests from the Kit of Factor Referenced Cognitive Tests: controlled associates, opposites and figures of speech); verbal episodic memory (immediate recall from two word list learning and two story memory tasks).	After age at baseline and gender were controlled, higher education was related to better performance in all cognitive domains, especially verbal fluency. The effect of education was the smallest for the processing speed domain. However, higher education was not associated with reduced rate of decline in any cognitive domain. Considering education as a dichotomous variable did not alter this pattern of results. Excluding the covariate of baseline age and running separate models in subgroups of younger (<70 years) and older (>70 years) still revealed no association between education and the trajectory of cognitive decline.

APOE, apolipoprotein; BNT, Boston Naming Test; CIQ, crystallized intelligence quotient; MMSE, Mini-Mental State Examination; NA, not available; NART, National Adult Reading Test; RAVLT, Rey Auditory Verbal Learning Test; SMLT, Symbol-Letter Modalities Test.

Example

- Lenehan identified multiple reasons that could explain different results
- Amongst them, differences in *statistical analysis* used appeared as a significant source of heterogeneity of results
- The *coordinated analytical approach* proposed by Piccinin et al (2013) aims at reducing differences due to the fit of different statistical models.
- Same model fitted independently to various studies
- Consistent data coding (continuous covariates centered at same values; categorical variables also coded consistently)
- Sensitivity analyses are essential
- Coordinated approach will facilitate the comparison of results & identification or patterns of results

Coordinated analysis approach (Piccinin et al, 2013)

- MMSE scores from 6 longitudinal studies of ageing:

-Canberra Longitudinal Study (CLS, English)

-Gerontological & Geriatric Population Studies of Gothenburg (H-70, Swedish)

-Healthy Older Person Edinburgh (HOPE, English)

-Origins of Variance in the Oldest Old (OCTO, Swedish)

-Longitudinal Aging Study of Amsterdam (LASA, Dutch)

-Swedish Adoption/Twin Study of Ageing (SATSA, Swedish)

- Most studies initiated in the early 1990s, except SATSA (1984) & H70 (1971, but MMSE first collected 1986)
- CLS (4 occ., ≈3.5 yrs apart); HOPE (4 occ., ≈4yrs apart);OCTO (5 occ., ≈2 yrs apart); LASA (5 occ., ≈3 yrs apart); SATSA (5 occ., ≈3 yrs apart)

Coordinated analysis approach

- LGM (regular & Tobit) fitted using time in study as time metric, with intercept and rate of change adjusted for age, sex (male=0, female=1) & education.

- Two sets of models were fitted:
  - Covariates centered at study specific median value:

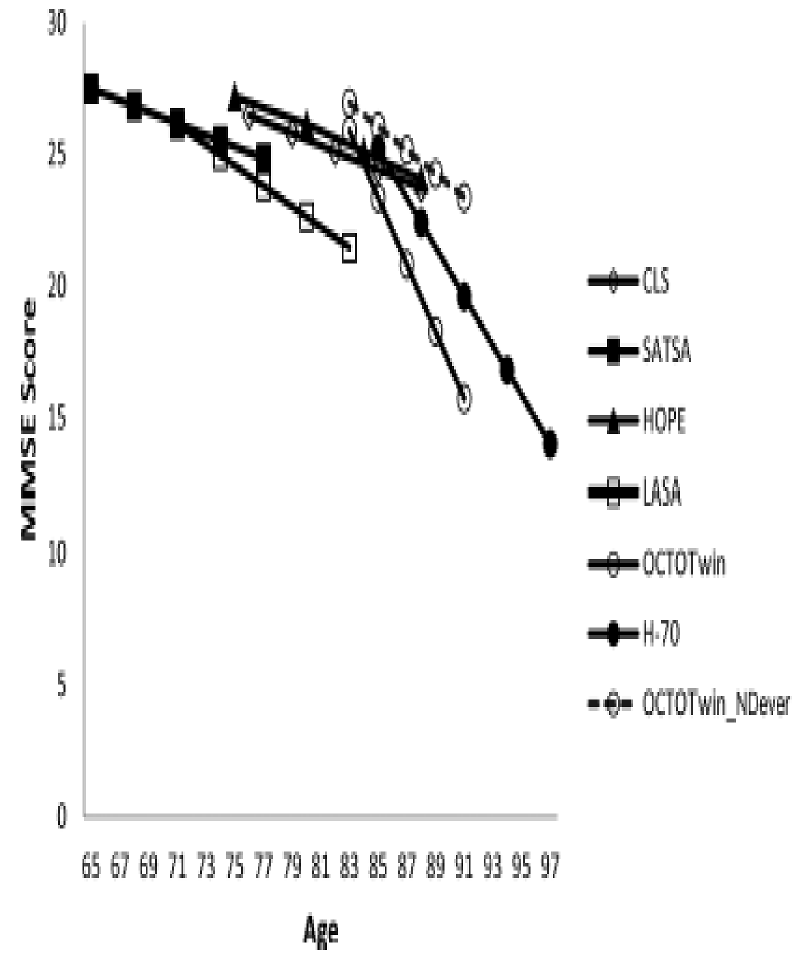
    Intercept & slope interpreted as expected value for men at median age with median level of education for the sample
  - Covariates centered at age 83 yrs and 7 yrs of education for common centering (except H70 & SATSA):

    Intercept & slope interpreted as expected values for men aged 83 with 7 yrs of education

H70 single initial age but aged 85 at 1st MMSE measurement ;H70 education already coded as <6 vs. >6 and SATSA with 4 categories

Coordinated analysis approach

- In all 6 studies, MMSE performance positively associated with level of educational attainment, controlling for age & sex
- In general, no association between education & rate of change except for OCTO
- Older individuals tended to score lower & declined at a faster rate



Conclusion

- Common centering: results essentially equivalent (except for intercept and linear slope means)
- Intercepts went down a bit (CLS: 24.20 vs 27.07 ; HOPE: 26.14 vs 27.84; LASA: 25.25 vs. 27.42; OCTO: 26.38 vs. 25.87)
- Rate of change moved towards OCTO's (CLS -0.42 vs. -0.22; HOPE -0.26 vs. -0.15; LASA -0.39 vs. -0.19; OCTO -1.20 vs. -1.27)
- *So, overall, different centering of covariates influenced trajectory parameter estimates but not their associations with the covariates*

Conclusion

- Coordinated analytical approach allowed us:
  - to perform a fairer comparison of results across studies
  - gain opportunities to understand reasons for diverse results
- As seen, may not be possible in all studies, but sometimes partial comparisons are still possible
- Provide best possible input for meta analysis of aggregate results if evidence synthesis is of interest.

References & suggested readings

- Griffith L. et al. Harmonization of cognitive measures in individual participant data and aggregate data meta-analysis. Methods Research Report.
- Piccinin et al. Coordinated analysis of age, sex and education on change in MMSE scores.J Gerontol B Psychol Sci Soc Sci. 2013 May;68(3):374-90.
- Hofer SM, Piccinin AM. Integrative Data Analysis through Coordination of Measurement and Analysis Protocol across Independent Longitudinal Studies. *Psychological methods*. 2009;14(2):150-164. doi:10.1037/a0015566.
- Bainter & Curran. Advantages of Integrative Data Analysis for Developmental Research.Journal of Cognition and Development , vol 16(1), 2015.
- Bauer & Hussong. Psychometric Approaches for Developing Commensurate Measures Across Independent Studies: Traditional and New Models. Psychological Methods, 2009, Jun; 14(2).
- Fortier et al. Quality, quantity and harmony: the DataSHaPER approach to integrating data across bioclinical studies. Int J Epidemiol. 2010 Oct;39(5):1383-93. doi: 10.1093/ije/dyq139. Epub 2010 Sep 2.