Data harmonization and individual patient data meta-analysis: exploring new avenues for evidence summaries and intervention development

Jovana Stojanovic, PhD, MSc, MPharm
Postdoctoral fellow
Department of Health, Kinesiology, and Applied Physiology;
Concordia University, Montreal, Canada

May 28, 2020

Contact: jovana.stojanovic@mail.concordia.ca
CONTENT

• Introduction

• Overview of the Action study (Decreasing sedentary behaviour and Increasing physical activity for healthy ageing)

Data harmonization: process & practical aspects

Individual Patient data (IPD) meta-analysis: process & practical aspects

• Conclusions
INTRODUCTION - POPULATION AGEING

- most advanced in Europe and in Northern America
- global population aged 60 and over:
  - 1980: 382 million
  - 2017: 962 million
  - 2050: 2.1 billion

79% of older individuals will be living in the developing regions

- GBD (195 countries 1990 – 2017): age-related disease burden accounts for 51.3% of all burden
• % of older adults meeting current guidelines: 2.4-83%
  
  ↓ overall mortality  
  ↓ CVD risk  
  ↓ breast and prostate cancer  
  ↓ fractures, and recurrent falls  
  ↑ functional and cognitive capacity  
  ↓ depression  
  ↓ dementia and Alzheimer

• $\approx 9.4 \text{ hr/day}$ (Objective)  
• $\approx 5.3 \text{ hr/day}$ sitting $\approx 3.3 \text{ hr/day}$ TV watching (Subjective)
  
  ↓ overall mortality  
  ↓ CVD risk  
  ↓ functional and cognitive capacity  
  ↓ autonomy  
  ↑ functional and cognitive decline  
  ↓ depression  

Lack of longitudinal evidence  
The importance of context $\rightarrow$ cognitively engaging SB
HEALTHY AGEING

healthy ageing = active ageing = successful ageing = productive ageing

Rowe and Kahn

- 1987
  - Usual ageing vs Successful aging

- 1997:
  - low probability of disease and disease-related disability and related risk factors
  - high cognitive and physical functional capacity
  - sustained engagement in social and productive activities

- 2015:
  - societal-level principles (opportunities for employment, voluntary work, social activity...)

Adapted from: Lu et al. Gerontologist, 2019;
HEALTHY AGEING

Number of papers measuring each domain of HA

- Physical capabilities: 37
- Cognitive function: 33
- Metabolic and physiological health: 26
- Psychological well-being: 24
- Social well-being: 22
- General health status: 16
- Security: 5
- Health behaviours: 3
- Short form survey: 6
- Health indices: 9

Adapted from: Lu et al. Gerontologist, 2019;
ACTION STUDY OVERVIEW

Through the use of existing international longitudinal studies, we aim to understand whether engaging in PA and SB after the age of 65 might contribute to healthy ageing in later life.
Screening for potentially eligible cohorts

- PUBMED (171 retrieves)
- Inter-university Consortium for Political and Social Research (77 studies)

### Studies

<table>
<thead>
<tr>
<th>Studies</th>
<th>Continent</th>
<th>Number of studies</th>
<th>Number of individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Americans' Changing Lives (ACL); Canadian longitudinal study on Ageing (CLSA); Cardiovascular Health Study (CHS); Charleston Heart Study; Cornell Study of Occupational Retirement; Costa Rican Longevity and Healthy Aging Study (CRELES); Framingham cohort study; Framingham Gen III; Framingham Offspring; Hispanic Established Population for the Epidemiological Study of the Elderly (HEPESE); Honolulu Heart Program (HHP); Long Beach Longitudinal Study; Midlife in the United States (MIDUS) Series; Multi-Ethnic Study of Atherosclerosis (MESA); National Health and Aging Trends Study (NHATS); National Social Life, Health, and Aging Project (NSHAP); Sacramento Area Latino Study on Aging (SALSA Study); Study of Women’s Health Across the Nation (SWAN) Series; The Health and Retirement Study (HRS); Women’s Health Initiative Study (WHI)</td>
<td>America</td>
<td>20</td>
<td>240K</td>
</tr>
<tr>
<td>Chinese Longitudinal Healthy Longevity Survey (CLHLS); Survey of Midlife in Japan (MIDJA); Social Environment and Biomarkers of Aging Study (SEBAS) in Taiwan</td>
<td>Asia</td>
<td>3</td>
<td>99K</td>
</tr>
<tr>
<td>The Irish Longitudinal Study on Ageing (TILDA); Swedish Adoption/Twin Study on Aging (SATSA); English Longitudinal study on Ageing (ELSA)</td>
<td>Europe</td>
<td>3</td>
<td>20K</td>
</tr>
<tr>
<td>The Survey of Health, Ageing and Retirement in Europe (SHARE); Study on Global AGEing and Adult Health (SAGE)</td>
<td>International</td>
<td>2</td>
<td>180K</td>
</tr>
<tr>
<td>Australian [Adelaide] Longitudinal Study of Aging</td>
<td>Australia</td>
<td>1</td>
<td>2K</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>29</strong></td>
<td><strong>540-545K</strong></td>
</tr>
</tbody>
</table>
Data harmonization

Process of retrospective data harmonization involves pooling heterogeneous data from disparate data sets and transforming them into 1 cohesive data set

Harmonized Dataset

Rolland et al., 2015;
Assemble information and select studies

Define variables and evaluate harmonization potential

Process data

Estimate quality of the harmonized dataset

Pre-harmonization
Project Proposal

Core process

Post-harmonization
Dissemination

Step 1

Step 2

Step 3

Step 4

Step 5

Step O

Data harmonization – simple example

**Study 1**
- Did you ever drink alcohol? 0=No 1=Yes
- If yes
  - Do you consume alcohol in the present time? 0=No, ex-drinker 1=Yes

**Harmonized variable:**
- Ever alcohol drinking 0=No 1=Yes

**Study 2**
- Do you currently drink alcohol? 0=No 1=Yes, less than 3 servings per week 2=Yes, more than 3 servings per week
- If no
  - Did you drink alcohol in the past? 0=No, never 1=Yes, more than 3 servings per week 2=Yes, less than 3 servings per week
<table>
<thead>
<tr>
<th>Measure</th>
<th>HRS</th>
<th>SHARE</th>
<th>TILDA</th>
<th>CRELES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Exercise</td>
<td><strong>Vigorous</strong></td>
<td>w1 - w13</td>
<td>w1-w2, w4-w7</td>
<td>w1-w2</td>
</tr>
<tr>
<td><strong>Activity</strong></td>
<td>aerobics, running, swimming, bicycling</td>
<td>sports, heavy housework, job involves physical labor</td>
<td>heavy lifting, digging, aerobics, fast bicycling</td>
<td>sports, jogging, dancing, or heavy work</td>
</tr>
<tr>
<td>Number of days/week</td>
<td>every day, more than 1 time/week, 1 time/week, 1-3 times/month, never</td>
<td>&gt; than 1 time/week, 1 time/week, 1-3 times/month, never</td>
<td>Number of days/week &amp; Number of hours/day</td>
<td>3 times/week</td>
</tr>
<tr>
<td>Unit</td>
<td>categorical</td>
<td>categorical</td>
<td>continuous</td>
<td>binary</td>
</tr>
<tr>
<td>Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Harmonized Vigorous physical exercise

- **0**: Never
- **1**: ≤ than 1 time/week
- **2**: ≥ than 1 time/week
- **3**: Everyday
WHAT IS IPD meta-analysis

- Instead of extracting aggregate data from the published literature, IPD meta-analysis allows the original participant level data to be requested and re-analysed
- ‘gold-standard’ in evidence summaries
Data harmonization and IPD meta-analysis in the era of open data

Number of SRs and MAs published from 1994 to 2014

Number of IPD-MAs published from 1994 to 2014

*Adapted from: Nevitt SJ et al. BMJ 2017; Ioannidis. Milbank Q. 2016;*
IPD meta-analysis methods: 2-STAGE APPROACH

Study 1
Study 2
Study 3
Study N

Study specific estimates

Standard fixed or random effects MA
IPD meta-analysis methods: 1 STAGE APPROACH
Multilevel model

- Inconsistently defined exposures and outcomes
- Complex data types, such as long-term outcomes, and time-dependent data
- Subgroup analysis and exploration of interactions
IPD vs. Aggregate data (AD) meta-analysis

- A review of 39 studies that compared IPD and AD meta-analysis of randomized trials

- 80% agreement in statistical significance
- ! 23 comparisons disagreed in direction of effect

- 20% disagreement in statistical significance
- ! 15% IPD-MA significant & AD-MA non-significant
- ! 5% IPD-MA non-significant & AD-MA significant

- 144 comparisons for the main effect
- 46 comparisons for effect modifiers

Tudur et al, Cochrane Database of Systematic Reviews 2016.
Data harmonization and IPD meta-analysis in the era of open data

Harmonization:

• https://g2aging.org/
• https://www.maelstrom-research.org/

Individual patient data access:

• https://biolincc.nhlbi.nih.gov/home/
• https://www.icpsr.umich.edu/web/pages/ICPSR/index.html
• https://vivli.org/
• https://yoda.yale.edu/
• https://clinicalstudydatarequest.com/Default.aspx
Opportunities

- Include unpublished, and/or poorly reported data
- Include studies with lots of missing data
- Standardize/harmonize outcomes and exposures
- Suitable for complex data types
- ↑ data quality
- ↑ generalizability of results
- ↑ statistical power
- Refined subgroup and sensitivity analysis, explore interactions
- Ability to focus on rare outcomes
- Encouraging more efficient secondary usage of existing data
- Collaboration
Challenges

• Time & resource-consuming
• Multi-disciplinary teams (data management and statistical expertise required)
• Heterogeneity of the collected variables
• Understanding the meaning of specific variables
• Ethical, legal, and consent-related restrictions (especially for international research):
  • Data sharing agreements
  • Confidentiality & protection agreements
• Availability (selection) bias
Performing an IPD meta-analysis: practical considerations

1. Research objectives:
   - Outcomes & exposures
   - Subgroups or covariates
   - Data type and modelling

2. Data availability
   - Adopt a systematic approach: search for all relevant published and unpublished studies → contact study authors
   - Convenience: include research groups from collaborative initiatives
     *Beware of availability bias: sensitivity with AD and IPD data*
     *PRISMA-IPD checklist*

3. Available resources

Stewart et al. JAMA.2015;
ACTION study → consideration for PA/SB intervention development in the ageing population

- development of evidence-based behavioral goals → improve health outcomes
- include environmental, organizational and societal perspective → WHO definition of an age-friendly community
- effective communication and messaging

ACKNOWLEDGEMENTS

ACTION study investigators:
International Research team
lead by Drs. Bacon and Lavoie

Simon L. Bacon, PhD
Professor
Concordia University, Canada
MBMC Co-director

Kim L. Lavoie, PhD
Professor
UQAM, Canada
MBMC Co-director

Funding bodies

MBMC Team

@ibtnetwork #ibtn2020
References: Healthy ageing

- Lu et al. Domains and Measurements of Healthy Aging in Epidemiological Studies: A Review. The Gerontologist, Volume 59, Issue 4, August 2019, Pages e294–e310,
References: IPD meta-analysis and data harmonization

• Nevitt et al. Exploring changes over time and characteristics associated with data retrieval across individual participant data meta-analyses: systematic review. BMJ 2017; 357: j1390.
• Ioannidis. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. The Milbank Quarterly, Vol. 94, No. 3, 2016 (pp. 485-514)
• Thomas et al. Systematic review of methods for individual patient data meta-analysis with binary outcomes. BMC Medical Research Methodology 2014, 14:79
• Ahmed et al. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. BMJ 2011; 344:d7762
• Riley et al. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ 2010; 340:c221