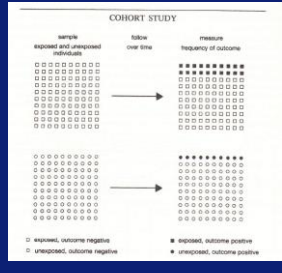
Statistical analysis of longitudinal studies: how to evaluate changes in risk factor and outcome?

Alex Burdorf
Department of Public Health
Erasmus MC Rotterdam

Disclosure

This presentation expresses my views, shaped by discussions with co-authors, colleagues, conference participants, and (unknown) reviewers of manuscripts

The longitudinal study design: the classical definition



Exposure:
* at start of enrolment
* during follow-up period
* frequency, duration, level

Time:
* years at risk
* time until event (latency period)
* relevant time window

Outcome:
* incidence of disease & mortality
* occurrence of consequence

Last, 2001

Changes in risk factors and health outcomes over time

Patterns of change:

1. **Dynamic pattern of health outcome:**
   e.g. common diseases such as musculoskeletal disorders
2. **Time-varying risk factors:**
   e.g. working conditions
3. **Dynamic interrelation between risk factor and health outcome that changes over time**
   e.g. chronic diseases and paid employment

Time-varying outcome measure

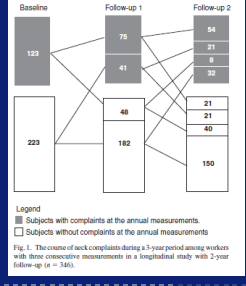


Fig. 1. The course of neck complaints during a 3-year period among workers with three consecutive measurements in a longitudinal study with 2-year follow-up (n = 346).

Measures of health outcome:

- **Prevalence ?**
- **Incidence ?**
- **Recurrence ?**
- **Recovery ?**

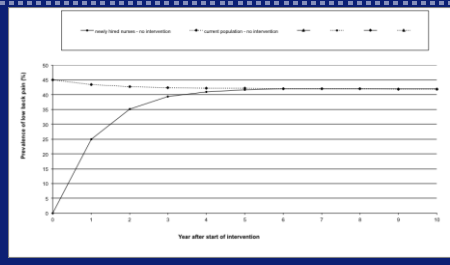
Luime et al. J Clin Epidemiol 2005;58:407-13

Time-varying outcome measure

	Neck		
	Baseline	Follow-up 1	Follow-up 2
Complaints in the past 12 mo, % (95% CI)			
Prevalence	35.6 (30.6–40.7)	33.5 (28.5–38.5)	33.2 (28.2–38.2)
Incidence		18.4 (13.3–23.5)	17.4 (12.5–22.3)
Recurrence		61.0 (54.6–67.4)	64.7 (58.5–70.9)
Complaints present in at least 3 of the past 12 mo, % (95% CI)			
Prevalence	8.7 (5.7–11.7)	8.7 (5.7–11.7)	10.1 (6.9–13.3)
Incidence		6.6 (3.3–9.9%)	5.7 (2.7–8.7)
Recurrence		30.0 (22.0–36.0)	56.7 (50.3–63.1)

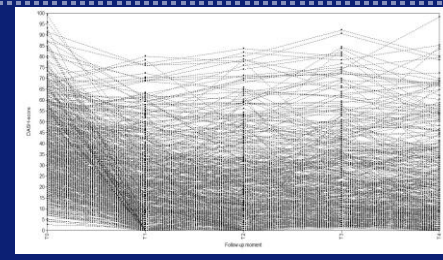
Luime et al. J Clin Epidemiol 2005;58:407-13

Time-varying outcome measure (Markov multistate model)



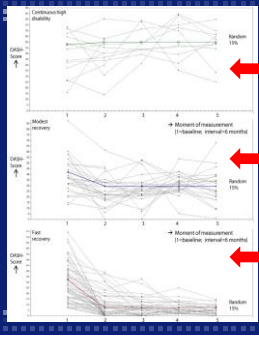
Population measures: incidence 25%, recurrence 66%, prevalence 45%
Markov model: epidemiologic measures converted into transition probabilities
Burdorf et al, Occup Environ Med 2013;70:491-7

Time-varying outcome measure (trajectories)



Prognosis of non-specific complaints of arm, neck, and shoulder among subjects at GP
Miedema et al. Physical Therapy 2016;96:972-84

Time-varying outcome measure
(latent class growth mixture modeling)



Three clinical patterns of limitations:

Continuous high level of limitations (8.8%)

Modest recovery over time (23.6%)

Fast recovery of time (67.6%)

Individual predictability ?

Time-varying determinant and outcome measure

Summary, sofar:

1. **Many diseases and disorders have a dynamic pattern over time**
2. **For consequences of disease also a highly dynamic pattern can be expected**

The same applies to many risk factors:
→ **Dynamic interrelation between risk factor and health outcome that changes over time**

Examples from two PhD students to illustrate this interrelation

Time-varying determinant and outcome measure

Example 1

Table 2: Adjusted association between health behaviors, work-related characteristics, and work engagement with work ability and sickness absence among employees (n=733).

	Less than good work ability (n=65) OR (95% CI)	1-9 sickness absence days (n=320) OR (95% CI)	10 or more sickness absence days (n=67) OR (95% CI)
Health behavior			
Body Mass Index			
Normal weight	1.00	1.00	1.00
Overweight	0.93 (0.51-1.70)	0.96 (0.68-1.35)	1.83 (0.97-3.45)
Obese	1.12 (0.51-2.46)	0.60 (0.34-1.04)	2.44 (1.12-5.35)*
Work-related characteristics			
High work demands	2.23 (1.24-3.99)*	1.00 (0.70-1.44)	1.14 (0.62-2.13)
Low skill discretion	2.19 (1.23-3.90)*	1.35 (0.92-1.99)	1.53 (0.81-2.88)
Work engagement			
Low work engagement	3.68 (2.15-6.30)*	1.36 (0.95-1.95)	1.84 (1.04-3.27)*

The Contribution of Work Engagement to Self-Perceived Health, Work Ability, and Sickness Absence Beyond Health Behaviors and Work-Related Factors

Anne Rongen, MSc, Suzan J. W. Robroek, PhD, Wilmar Schaufeli, PhD, and Alex Burdorf, PhD

JOEM • Volume 56, Number 8, August 2014

6 months follow-up

Time-varying determinant and outcome measure

Example 1

Table 2: Adjusted association between health behaviors, work-related characteristics, and work engagement with work ability and sickness absence among employees (n=733).

	Less than good work ability (n=65) OR (95% CI)	1-9 sickness absence days (n=320) OR (95% CI)	10 or more sickness absence days (n=67) OR (95% CI)
Health behavior			
Body Mass Index			
Normal weight	1.00	1.00	1.00
Overweight	0.93 (0.51-1.70)	0.96 (0.68-1.35)	1.83 (0.97-3.45)
Obese	1.12 (0.51-2.46)	0.60 (0.34-1.04)	2.44 (1.12-5.35)*
Work-related characteristics			
High work demands	2.23 (1.24-3.99)*	1.00 (0.70-1.44)	1.14 (0.62-2.13)
Low skill discretion	2.19 (1.23-3.90)*	1.35 (0.92-1.99)	1.53 (0.81-2.88)
Work engagement			
Low work engagement	3.68 (2.15-6.30)*	1.36 (0.95-1.95)	1.84 (1.04-3.27)*

Referent: "the statistical analysis did not take into account all data available (…) no adjustment for baseline outcomes (…) inappropriate analysis"

What to do ?

Time-varying determinant and outcome measure

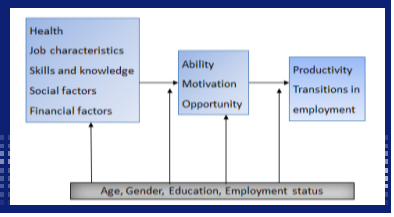
Example 2

Discussion on study about influence of chronic disease on work ability:

1. **How to describe dynamic pattern of chronic disease (= determinant) and work ability (= outcome) over time ?**
2. **Influence of chronic disease on work ability and of work ability on chronic disease ?**
3. **How to deal with repeated observations ?**

STREAM: Study on Transitions in Employment, Ability and Motivation

- Aged 45-64 years at baseline
- Employees, self-employed and non-employed
- Online questionnaire study



Time-varying determinant and outcome measure

Example 2

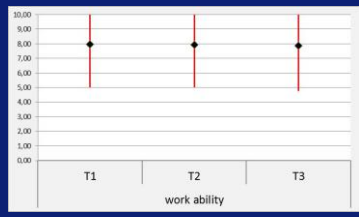
Basic question:
How do chronic diseases influence work ability and productivity at work ?

STREAM cohort

- **study population:** n = 8411 persons, aged 45 plus, with paid employment
- **study design:** annual questionnaires (3x)
- **determinant = chronic disease:**
  "*Do you [currently] have one or more of the following chronic diseases, disorders, or handicaps?***"** [yes/no]
- **outcome = work ability:**
  "*How would you rate your work ability at this moment relative to the best time of your life?***"** [0 = not able to work to 10 = work ability at lifetime best]

Variance in outcome measure (work ability)

Example 2



Questions:
- how can chronic disease explain within worker variance (37%) ?
- how can chronic disease explain between worker variance (63%) ?

Variance in chronic diseases (determinant)

Example 2

	Musculoskeletal complaints		Psychological complaints	
Baseline T1 prevalence	32%		3.6%	
Follow-up	T2	T3	T2	T3
prevalence	33%	34%	3.7%	3.4%
'incidence'	15%	14%	1.9%	1.7%
'recurrent'	71%	73%	51%	48%
'recovery'	29%	27%	49%	52%

Chronic disease and work ability

Erasmus MC Ezafung

Example 2

Observations so far in STREAM:

1. **Chronic diseases and work ability are measures of occurrence (level)**
2. **Chronic disease and work ability show large variation over time**
3. **Large variation between persons suggests systematic difference between persons in the study (covariates, chronic disease?)**
4. **Large variation within persons suggests that determinant and outcome may vary simultaneously (analysis of change)**

Chronic disease and work ability (statistical models)

Erasmus MC Ezafung

Example 2

Table 1. Three specifications of a Generalized Estimating Equations (GEE) model in the analysis of the influence of health problems on work ability and productivity at work in a longitudinal study with three annual waves (T1, T2, and T3). Note: The same analyses were done with productivity at work as the outcome.

Models	Outcome (Y)	Predictors (X)	Covariates
Time-lag $Y_{(t)} = \beta_0 + \beta_1 X_{(t-1)}$	Work ability (T2)	Health problem (T1)	Other health problems (T1) Work-related factors (T1) Individual factors (T1)
	Work ability (T3)	Health problem (T2)	Other health problems (T2) Work-related factors (T2) Individual factors (T1)
Autoregressive $Y_{(t)} = \beta_0 + \beta_1 X_{(t-1)} + \beta_2 Y_{(t-1)}$	Work ability (T2)	Health problem (T1)	Other health problems (T1) Work-related factors (T1) Individual factors (T1) Work ability (T1)
	Work ability (T3)	Health problem (T2)	Other health problems (T2) Work-related factors (T2) Individual factors (T1) Work ability (T2)
Incidence-recovery $Y_{(t)} = \beta_0 + \beta_1 (X_t . X_{t-1}) + \beta_2 Y_{(t-1)}$	Work ability (T2)	Incident health problem (T1–T2) Recovered health problem(T1–T2)	Other health problems (T1) Work-related factors (T1) Individual factors (T1) Work ability (T1)
	Work ability (T3)	Incident health problem (T2–T3) Recovered health problem(T2–T3)	Other health problems (T2) Work-related factors (T2) Individual factors (T1) Work ability (T2)

Leijten et al. Scand J Work Environ Health 2014;40:473-82.

Chronic disease and work ability

Erasmus MC Ezafung

Example 2

Cross-sectional analysis

Chronic disease T1 – work ability T1

	Work ability (0-10 scale)	
	β	se
Musculoskeletal complaints	-0.42	0.03
Psychological complaints	-1.55	0.08

Chronic disease and work ability

Erasmus MC Ezafung

Example 2

Longitudinal analysis (time lag model)

Chronic disease T1 → work ability T2

	T1 → T2 Work ability		T1 – T1 Work ability	
	β	se	β	se
Musculoskeletal complaints	-0.37	0.03	-0.42	0.03
Psychological complaints	-1.02	0.08	-1.55	0.08

Chronic disease and work ability

Erasmus MC Ezafung

Example 2

Cross-sectional versus longitudinal analysis:

Interpretation:

- **presence of chronic disease is associated with lower work ability**
- **longitudinal analysis is very similar to cross-sectional analysis**

Potential solution:

- **since work ability at baseline will strongly predict value at follow-up, *and* disease and work ability are measured in a questionnaire, adjustment for baseline value is often advised to get rid of 'common source bias'**

Chronic disease and work ability

Erasmus MC Ezafung

Example 2

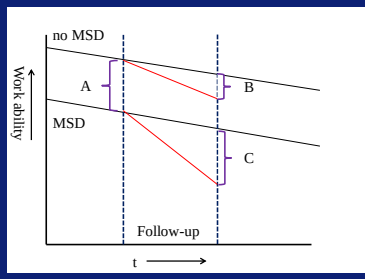
Longitudinal analysis (autoregressive model)

Chronic disease T1 → work ability T2, adjusted for work ability T1 (presence of chronic disease on change in outcome)

	T1 → T2/T1 Work ability		T1 → T2 Work ability	
	β	se	β	se
Musculoskeletal complaints	-0.15	0.02	-0.37	0.03
Psychological complaints	-0.30	0.05	-1.02	0.08

Time-varying determinant and outcome measure

Erasmus MC Ezafung
Example 2



Two hypotheses:

1. Traditional analysis
Systematic difference in work ability between healthy workers and those with MSD (= A)

2. Analysis of change
Decrease in work ability is faster among workers with MSD than those healthy (= C – B)

Chronic disease and work ability

Erasmus MC Ezafung
Example 2

Is adjustment for baseline value of work ability really the solution for common source bias?

The adjustment will answer a different question!

Kristensen et al. SJWEH 2013:
- **driving force of the analysis should be the hypothesis**
- **we need insight of how changes in determinants influence changes in outcome**

Solution:
- **analyse how change in presence of chronic disease will influence work ability**

Chronic disease and work ability

Erasmus MC Ezafung
Example 2

Longitudinal analysis (incidence-recovery model)

Onset or recovery chronic disease during T1 – T2, against work ability T2, adjusted for work ability T1 (Change in chronic disease on change in outcome). Analysis stratified by disease status at baseline

			T1/T2 → T1/T2 Work ability β	se	T1 → T2 Work ability β	se
Musculoskeletal complaints	-	onset	-0.30	0.04	-0.37	0.03
	-	recovery	+0.09	0.05		
Psychological complaints	-	onset	-1.50	0.08	-1.02	0.08
	-	recovery	+0.67	0.15		

Chronic disease and work ability

Erasmus MC Ezafung
Example 2

Interpretation:

- **separate the direction of change (onset vs recovery)**
- **onset has stronger effect than reciprocal effect of recovery:**
  - *** insufficient recovery ?**
  - *** disease history before enrolment in the cohort ?**
- **responsiveness of influence of chronic disease on work ability:**
  - *** relevant time window shorter than one year ?**

Chronic disease and work ability

Erasmus MC Ezafung
Example 2

Original article

Scand J Work Environ Health 2014;40(5):473-482
doi:10.5271/sjweh.3444

The influence of chronic health problems on work ability and productivity at work: a longitudinal study among older employees

by Leijten FRM, van den Heuvel SG, Ybema JF, van der Beek AJ, Robroek SJW, Burdorf A

Detailed analyses and discussion on options in statistical analysis

Dynamic interplay of determinants, health, and consequences over time

Erasmus MC Ezafung

Further developments:

1. **Statistical models for analyzing change:**
   Generalized Estimating Equations (population average effect; example 2**)**
   Random coefficient and slope model (allowing for differences between subjects**)**

   Fixed effects models (only change within individuals**)**
   Hybrid regression models (separate predictors for within- and between subject variation**)**

2. **Models for change over life course:**
   Trajectory analysis
   Transition probabilities deterministic Markov model

Time-varying determinant and outcome measure

Advice:

1. **There is no such thing as THE best strategy for longitudinal analysis in all situations: be critical !**

2. **Analyses on occurrence or change in occurrence will answer completely different questions: specify your theory !**

3. **Crucial considerations:**
   - *** variability of determinant and outcome over time**
   - *** is the determinant a unique event in a given period or a reflection of historical pattern**
   - *** is the outcome a unique event in a given period, independent from other possible outcomes**
   - *** relevant time-window between cause and effect (responsiveness)**

a.burdorf @ erasmusmc.nl

Bram van der Leck, Workers in port, 1916