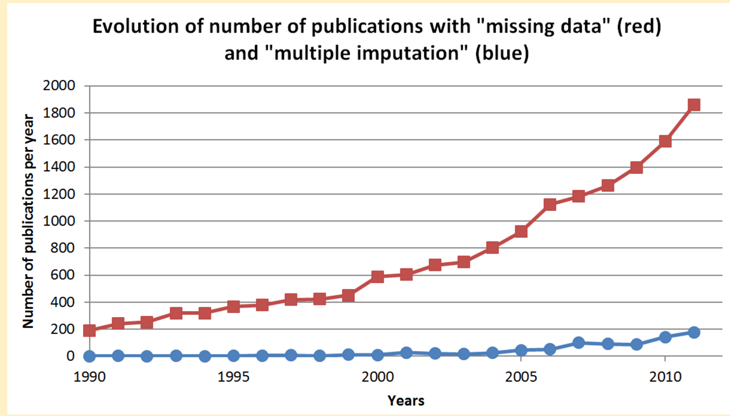
Interest of using multiple imputations. Assessment of work-related risk factors for the incidence of lateral epicondylitis

E Herquelot [1,2], J Bodin[3], Y Roquelaure[3], C Ha[4], A Leclerc [1,2], M Goldberg[1,2], M Zins[1,2], A Descatha[1,2,5]

(1) Versailles St-Quentin University, France
(2) Inserm U1018, Centre for research in Epidemiology and Population Health, Population-Based Epidemiological Cohorts Research Platform, France
(3) LUNAM University, Laboratory of Ergonomics and Epidemiology in Occupational Health, University of Angers, France
(4) Department of Occupational Health, French Institute for Public Health Surveillance (InVS), France
(5) AP-HP, Occupational Health Unit/EMS, University hospital of West suburb of Paris, France

CONTEXT

- Increasing **concern on missing data** with low participation rates in recent studies
  - ⇒ Estimations on available data could be strongly biased
  - ⇒ Problem of statistical methods to deal with missing data
- **Multiple imputations** = effective method to **incorporate the knowledge of investigators** and extract the **maximum potential of available data** (1)
  - ⇒ Multiple imputation can **reduce bias** of estimation on available data



OBJECTIVES

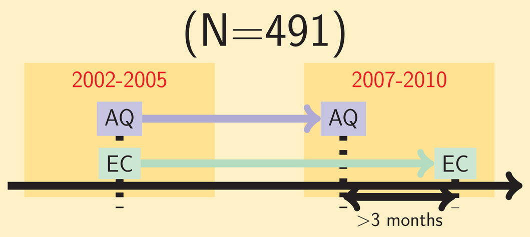
To analyze the **effects of occupational risk factors**, measured twice, on **incidence of lateral epicondylitis** using multiple imputations analysis.

ORIGINAL DATA

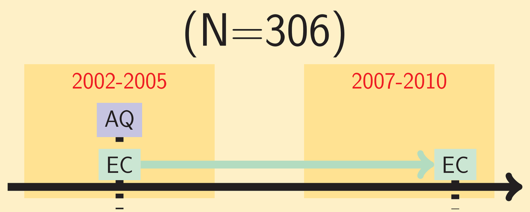
Study population

- Large sample of workers in the Loire Valley district of Central West France in **two successive surveys** in 2002-2005 and in 2007-2010 with self-assessment of **occupational exposures (AQ)** and **physical examination on lateral epicondylitis (EC)** by an occupational physician (2)
- 3710 workers included in 2002-2005
  - ⇒ Study population = 3231 workers without elbow pain or lateral epicondylitis in 2002-2005
- For the 1881 male workers:

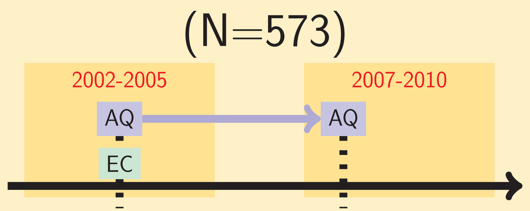
Complete case workers (N=491)



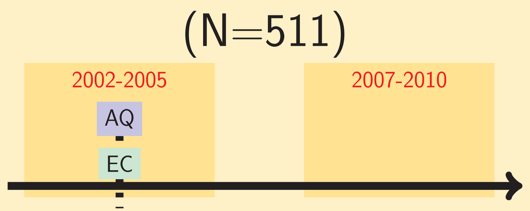
Physical examination only (N=306)



Self-reported questionnaire only (N=573)



Lost of follow-up (N=511)



Risk factors

Previously found on prevalence in the literature and in the same population (3):

- **Age** in 3 class (<30, 30-44, ≥ 45)
- **Repetitive Tasks** (>4 hours a day)
- **Specific elbow combined physical exposure**: high physical exertion associated with elbow flexion/extension or extreme wrist bending (>2 hours a day)

Missing variables at second survey (2007-2011)

- Outcome: lateral epicondylitis at physical examination
- Risks factors: time between the physical examinations (offset), high physical exertion, elbow flexion and extension, extreme wrist bending, high repetitiveness, low social support,
- Auxiliary variables: job change, self-assessment of elbow pain

MULTIPLE IMPUTATIONS

We used **Multiples Imputations by Chained Equation algorithm** (MICE) to impute the 9 missing variables using the following variables:

- in 2002-2005: year of first questionnaire, department, body mass index, occupations, socio-economic category, type of contract combined with working experience, lifting and carrying objects, past history of upper-extremity musculoskeletal disorders, rotator cuff syndrome, lateral epicondylitis
- in 2007-2010: professional change since 2002, the declaration of pain at elbow, having a second physical examination, lateral epicondylitis

Missing At Random hypothesis (MAR)

Missing data on 9 missing variables depended on previous measured variables.

Missing Not At Random hypothesis (MNAR)

Missing values on lateral epicondylitis depended on previous measured variables and on unmeasured variables at second examination:

- Health status
- Working condition

} approximate by cause of absence at second examination (No appointment, Change in situation, Not known)

- Employment instability } approximate by age

	First scenario	Second scenario
Chosen prevalence	by cause of absence	by cause of absence and age

STATISTICAL ANALYSIS

Poisson models were performed to assess the incidence rate ratios (IRRs) of risk factors separately by sex:

- Main analysis: analysis on complete-case workers and MAR imputed workers
- Complementary analysis: analysis on MNAR imputed workers

The results presented here are restricted to men.

MAIN ANALYSIS

Annual incidence rate of lateral epicondylitis = 1.0[0.7;1.3] per 100 men

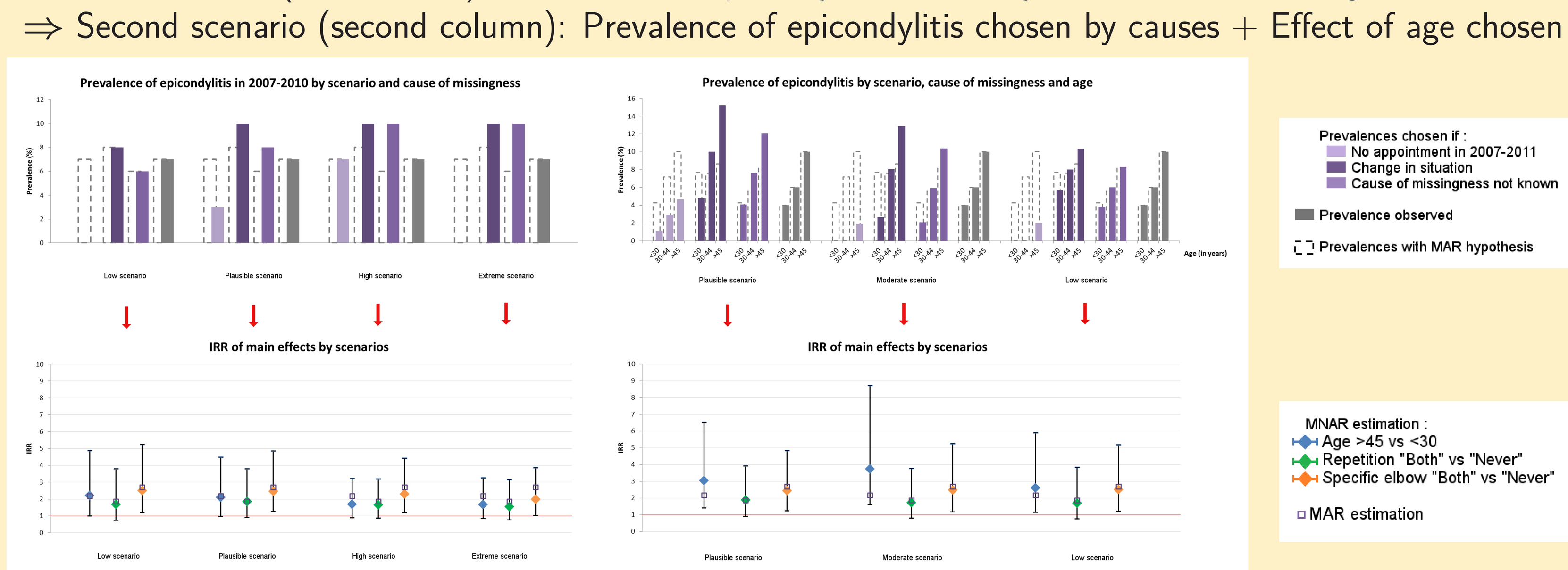
	Complete case, N=491				Multiple imputation (MI), N=1881			
	N	**N event**	**IRR (95% CI)**	**p**	**N**	**N event**	**IRR**	**p**
Age, in years								
<30	81	5	1.0	.	452	18.2	1.0	.
30-44	267	14	0.8 (0.3- 2.1)	0.69	857	44.8	1.4 (0.6- 3.5)	0.44
≥ 45	143	8	0.9 (0.3- 2.6)	0.86	572	40	2.2 (0.9- 5.4)	0.09
Doing repetitive tasks, more than 4 hours/day								
Never exposed	350	17	1.0	.	1297.9	58.6	1.0	.
Exposed at first questionnaire	45	2	1.1 (0.3- 4.5)	0.87	217.9	12.8	1.2 (0.4- 3.6)	0.78
Exposed at second questionnaire	54	2	0.5 (0.1- 2.1)	0.36	192.1	10.9	0.8 (0.3- 2.6)	0.77
Exposed at both questionnaires	42	6	**2.6 (1.0- 6.7)**	**0.05**	173.1	20.6	1.9 (0.8- 4.6)	0.17
Specific elbow combined physical exposure								
Never exposed	302	14	1.0	.	1087.9	41.0	1.0	.
Exposed at first questionnaire	*65*	*0*	*NE*		*324.7*	*13.8*	*NE*	
Exposed at second questionnaire	54	6	**2.5 (1.0- 6.5)**	**0.05**	201.1	18.7	2.5 (1.0- 6.0)	**0.05**
Exposed at both questionnaires	70	7	1.7 (0.6- 4.2)	0.29	267.3	29.5	**2.7 (1.2- 6.0)**	**0.01**

⇒ Age NS in both analysis, value correspond in MI analysis as what it is expected

⇒ Specific elbow exposure S for exposed at both questionnaires in MI analysis

COMPLEMENTARY ANALYSIS

- MAR hypothesis (dashed bars)
  - ⇒ Prevalence of epicondylitis similar between the three categories of causes of absence
  - ⇒ Effect of age identical between categories of causes of absence
- Sensitivity analyses with MNAR hypothesis (solid bars)
  - ⇒ First scenario (first column): Prevalence of epicondylitis chosen by causes + Effect of age identical
  - ⇒ Second scenario (second column): Prevalence of epicondylitis chosen by causes + Effect of age chosen



⇒ Sensitivity analyses indicated that the results were robust.

CONCLUSION

- **Interest of multiple imputations when the proportion of missing data is large**, in order to reduce the effect of attrition (potential bias)
- Checking robustness of results with **comparison with complete-case analyses and sensitivity analyses could be recommended**

References :
(1) White IR et al. Multiple imputation using chained equations: Issues and guidance for practice. Statistics in medicine. 2011
(2) Roquelaure Y et al. Epidemiologic surveillance of upper-extremity musculoskeletal disorders in the working population. Arthritis Rheum. 2006
(3) Herquelot E et al. Work-related risk factors for lateral epicondylitis and other cause of elbow pain in the working population. Am. J. Ind. Med. 2012

 eleonore.herquelot@inserm.fr