

# Data sharing: What about epidemiological cohorts?

Marcel Goldberg & Marie Zins

PIs of the GAZEL and CONSTANCES cohorts

Inserm Unit 1018, Villejuif, France

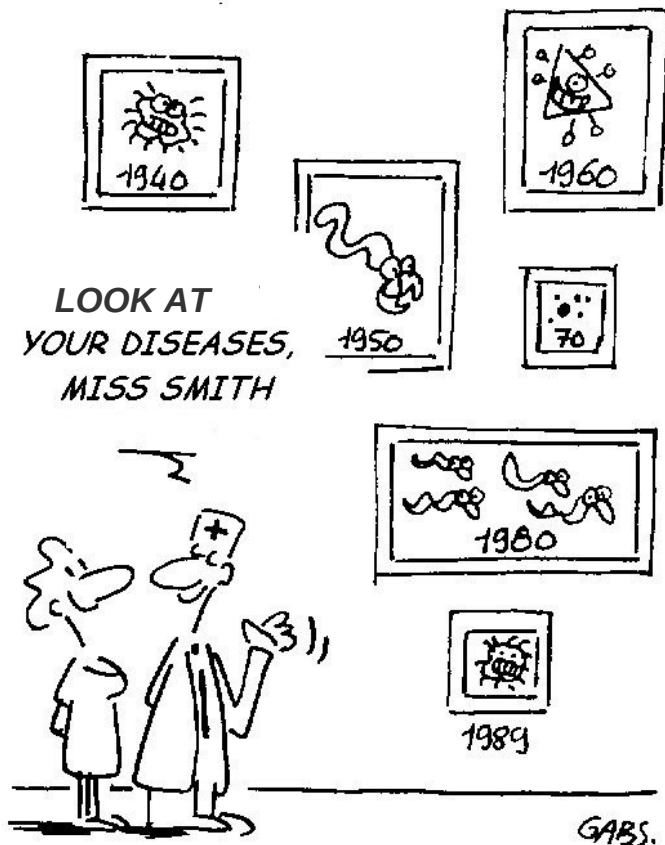
# Different types of epidemiological cohorts

- Cohorts of patients suffering from a specific disease
  - aimed at studying the course of the disease
  - usually “small” (hundreds or thousands of patients)
- Population-based cohorts
  - aimed at analysing the causes of and risk factors for different health outcomes
  - large-scale: dozens or hundreds of thousands subjects
  - incorporate personal, social, lifestyle, occupational and environmental data as well as biobanks of blood and other biological specimens
- “General-purpose cohorts”
  - shaped to cover a broad scope of determinants and outcomes
  - designed to answer several research questions, even if these questions were not defined at inception
  - constitute research platforms open to the research community for developing multiple nested projects

# The GAZEL cohort: an Open General-Purpose Epidemiological Laboratory

- Composition
  - 20,625 men (15,011) and women (5,614)
  - men aged 40-50; women 35-50
  - working at the French public utility company (EDF-GDF )
- Inception in 1989
- Long term follow up
  - systematic and prospective collection of various data from multiple sources
  - offers an easy access to the data base and to volunteers for additional data collection

# Epidemiological research in the cohort



- More than 40 specific research projects from 20 different research teams: France, Belgium, Canada, Finland, Germany, Sweden, UK, USA
- Health problems: CHD, cancer, depression, menopause & osteoporosis, migraine, musculoskeletal disorders, general state of health, mortality..
- Personal & Occupational risk factors: tobacco & alcohol, diet, psychosocial factors, life events, social support, occupational exposures, work organisation & constraints, traffic accidents...
- General & Social factors: social determinants of health, health inequalities

# Role of the GAZEL Team

- Global design of the project
- Design of the questionnaires, biobank, data from external databases...
- Follow-up of the participants
- Management of the GAZEL database (security, confidentiality...)
- Funding of the general functioning of the cohort

# Rules for nested research projects and access to the database

[http://www.gazel.inserm.fr/index.php?option=com\\_content&view=article&id=17&Itemid](http://www.gazel.inserm.fr/index.php?option=com_content&view=article&id=17&Itemid)

# Application for a new research project

- Every research group is entitled to apply
  - In France or in other countries
  - Public or private
- Formal application to the GAZEL team (objectives, description of sample, methods, data needed, timetable); the PI must have a permanent position in his/her organization
- If non anonymous data are needed, all legal and ethic authorizations must be provided
- Validation committee (Inserm)
  - Scientific quality
  - Conflicts of interest (including with existing research projects)
- For selected projects:
  - Only the PI of the project is entitled to ask for data (except specific arrangements)
  - A member of the Gazel team is chosen as the regular scientific correspondent of the project
- All specific research projects must find their own additional funding when necessary (collection of supplemental data, extraction of biospecimen)

# Types of data

1. Data routinely collected by the GAZEL team and integrated in the “open Gazel data base” (GAZEL database Catalog)
2. Data collected by other research groups for specific projects
3. Genetic and biological material



# Data routinely collected

- All anonymous data from the open Gazel data base can be provided
- A precise list of data must be established (definition of the sample, data list, period...)
- Modalities of transfer (media, date...) are established in collaboration between the PI and corresponding member of the Gazel team
- Analyses are allowed only accordingly to the protocol described in the application; in case of major changes a new application may be necessary
- Data can not be transferred to persons not mentioned in the application
- When the project is finished, the data can not be used any more

# Data collected by research groups other than the GAZEL team

- Data may be collected directly by the PI or through regular GAZEL questionnaires according to specific arrangements
- Data collection modalities must have been previously described in details in the application and all legal requirements fulfilled; only data listed in the application can be collected
- All correspondence with the cohort participants must be previously approved by the GAZEL team
- All data collected by other investigators are their own property and can be freely used by them
- For security reasons a copy of the files must be provided to the GAZEL team; these files are for archives only and will not be used or transferred to others without a written permission of the PI
- After a period defined in agreement with the GAZEL team, all data collected by other investigators can be integrated in the common GAZEL database along with a precise documentation allowing for their use by others

# Access by others to the data collected by a research group

- Sharing of data is encouraged
- Data already integrated in the common GAZEL database: same rules as for the open GAZEL database
- Data not integrated in the common GAZEL data base
  - Already published tools (scales, specific questionnaires...)
    - Data were collected through regular GAZEL questionnaires : GAZEL PIs decide whether they can be transferred
    - Data were collected directly by an external research group: same as original data (see below)
  - Original tools and data (designed and collected directly by an external research group): the PI responsible of the data collection decides whether he/she wants to share them; arrangements directly between researchers (GAZEL PIs must be informed)

# Genetic and biological material

## Special rules apply due to French regulation

- Specific application to the French official body
- Informed consent of the individuals
- Financial participation
- Rarity of the material

# Potential conflicts between researchers

- A new project may interfere with current ones
- GAZEL team PIs try to avoid conflicts by establishing a direct relationship between researchers
- If no agreement can be reached directly between researchers, GAZEL team PIs propose the solution that they consider the best for both parties
- If the researchers do not agree, the advice of ADELFF (the French speaking epidemiologists association) is asked
- If the researchers still do not agree, GAZEL team PIs make a final decision

# Publications

- Any kind of scientific publication (journal, congress, thesis...) is under the full responsibility of the PI of the research projects according to his/her own rules and no previous validation by GAZEL PIs is compulsory
- A copy of all published material (including reports and thesis) must be sent to the GAZEL team as soon as possible
- When asked, the PI of the research projects must prepare a summary of the main findings for a non-scientific audience for the GAZEL Journal, and must present them to the annual GAZEL Scientific symposium in Paris
- Signatures & Acknowledgments
  - The name of GAZEL must be included in the title whenever possible
  - At least one member of the GAZEL team (who has been a correspondent for scientific aspects and/or analyses) signs the publications; this is managed by the project PI and the GAZEL PIs
  - Acknowledgments must include GAZEL team

# Some difficulties with the recognition of the scientific work of the GAZEL team

## Refusal of a journal

*“Our colleague XX submitted for publication a manuscript on ... relying on data from the GAZEL Cohort Study. This study is managed by our group at INSERM from more than 20 years now, and is largely open to external researchers: they can freely benefit from the database provided that they follow some rules. Among these rules, we insist that the word “GAZEL” (or “GAZEL Cohort”, or “GAZEL Cohort Study”) is in the title of all papers that are published from our database. We think that is a common practice regarding the openness of large databases like GAZEL’s to the research community: I am sure that you understand that in return to give a free access to our data, we need to be easily identified in the publications, and when looking for publications from the “GAZEL Cohort Study”, one may use “GAZEL” in PubMed to search for all related publications.”*

# Publications

(in red: members of the GAZEL team)

- Articles **ON** the GAZEL cohort
  - **Goldberg M, Zins M et al.** Socioeconomic, demographic, occupational and health factors associated with participation in a long-term epidemiologic survey. A prospective study of the French GAZEL cohort and its target population. *Am J Epidemiol*, 2001, 154, 4, 373-384.
  - **Metzger MH, Goldberg M, Leclerc A, Zins M.** Factors associated with self-reporting of chronic health problems in the French GAZEL cohort. *J Clin Epidemiol*, 2002; 55: 48-59.
  - **Goldberg M, Zins M et al.** A. Health problems were the strongest predictors of attrition during follow up of the GAZEL cohort. *J Clin Epidemiol*. 2006;59:1213-21.
  - **Goldberg M, Zins M et al.** Cohort profile: The GAZEL Cohort Study. *Int J Epid*. 2007; 36:32-39 .
  - **Zins M, Leclerc A, Goldberg M.** The French GAZEL Cohort Study: 20 years of epidemiologic research. *Advances in Life Course Research* 2009; 14: 135-146.
- Recent Articles **FROM** the GAZEL cohort (out of 30 published in 2010-2011)
  - Silhol R, **Zins M**, Chauvin P, Chaix B. *J Epidemiol Community Health* **2011**
  - Sabia S, **Guéguen A**, Berr C, Berkman L, Ankri J, **Goldberg M, Zins M**, Singh-Manoux A. *Addiction* **2011**.
  - Stringhini S, Dugravot A, Shipley M, **Goldberg M, Zins M**, Kivimaki M, Marmot M, Sabia S, Singh-Manoux A. *PLoS Medicine* **2011**.
  - Lemogne C, Fossati P, Limosin F, Nabi H, Encrenaz G, **Bonenfant S**, Consoli S M. *Acta Psychiatrica Scandinavica* **2011**.
  - Sjösten N, Nabi H, Westerlund H, Singh-Manoux A, Dartigues JF, **Goldberg M, Zins M**, Oksanen T, Salo P, Pentti J, Kivimäki M, Vahtera J. *Cephalalgia* **2011**.
  - Buyck JF, Bonnaud S, Boumendil A, Andrieu S, **Bonenfant S, Goldberg M, Zins M**, Ankri J. *Am J Pub Health* **2011**.
  - Rod NH, Vahtera J, Westerlund H, Kivimaki M, Zins M, Goldberg M, Lange T. *Am. J. Epidemiol.* **2010**.
  - Westerlund H, Vahtera J, Ferrie JE, Singh-Manoux A, Pentti J, Melchior M, Leineweber C, Jokela M, Siegrist J, **Goldberg M, Zins M**, Kivimäki M. *BMJ* **2010**.
  - Fournier A, Kernaleguen C, Panjo H, Clavel-Chapelon F, Ringa V. *Menopause.* **2010**
  - Galéra C, Bouvard MP, Melchior M, Chastang JF, Lagarde E, Michel G, Encrenaz G, Messiah A, Fombonne E. *Eur Psychiatry.* **2010**.
  - Melchior M, Chastang JF, Walburg V, Arseneault L, Galéra C, Fombonne E. *Depress Anxiety.* **2010**.
  - Melchior M, Ferrie JE, Alexanderson K, **Goldberg M**, Kivimaki M, Singh-Manoux A, Vahtera J, Westerlund H, **Zins M**, Head J. *Am J Epid* **2010**.
  - Ringa V, Fritel X, Varnoux N, Zins M, Quelen C, Bouyer J. *Fertil Steril* **2010**.
  - Berr C, Vercambre MN, Bonenfant S, Singh Manoux A, **Zins M, Goldberg M.** *Dementia and Geriatric Cognitive Disorders* **2010**.
  - Dugravot A, Sabia S, Stringhini S, Kivimaki M, Westerlund H, Vahtera J, **Guéguen A, Zins M, Goldberg M**, Nabi H, Singh-Manoux A. *Am J Clin Nutrition*, **2010**.
  - Wahrendorf M, Ribet C, **Zins M, Goldberg M**, Siegrist J. *Ageing and Mental Health* **2010**.
  - Lemogne C, Nabi H, **Zins M**, Cordier S, Ducimetière P, Goldberg M, Consoli SM. *Psychother Psychosom* **2010**.
  - Melchior M, Chastang JF, Mackinnon D, Galéra C, Fombonne E. *Drug and Alcohol Dependence.* **2010**.
  - Constant A, Lafont S, Chiron M, **Zins M**, Lagarde E, Messiah A. *Addiction*, **2010**.
  - Vahtera J, Westerlund H, Ferrie J, Head J, Melchior M, Singh-Manoux A, **Zins M, Goldberg M**, Alexanderson K, Kivimäki M. *J Epidemiol Community Health* **2010**.