



---

## Developing a data linkage system to enable innovative research

### International conference

September 3-4, 2013

9 a.m. - 6 p.m.

Room A 300

---

#### Tuesday, 3 September 2013

8.30-9.00 Registration

9.00-9.10 *Welcome and opening speech by the WZB President*  
Prof. **Jutta Allmendinger**, WZB (Berlin)

#### *The Western Australia (WA) data linkage system and the Developmental Pathways in WA Children Project*

##### **First session: Development of the WA data linkage system and its expansion**

(Chairperson: Dr. Jianghong Li)

9.10-10.15 *Australian data linkage infrastructure: Building on the WA experience.*  
Dr. **Merran Smith**, Chief Executive, Population Health Research Network (Perth).

10.15-10.35 *Linking the WA health data system to social outcome data: Why, how and what came out of it?*  
Dr. **Jianghong Li**, Senior Researcher, WZB (Berlin).

10.35-11.15 *Technical and practical aspects of cross agency data linkage: The West Australian example.*  
Prof. **Nicholas de Klerk**, Head of Biostatistics and Genetic Epidemiology, Telethon Institute for Child Health Research, Centre for Child Health Research, the University of Western Australia (Perth).

11.15-11.45 Coffee break

11.45–13.00 1) *Linking data to build an evidence base: Developmental Pathways in WA Children Project.*  
Dr. **Rebecca Glauert**, Program Manager, Developmental Pathways Project, Telethon Institute for Child Health Research, Centre for Child Health Research, the University of Western Australia (Perth).

2) *Developmental Pathways Project: Challenges for the WA data linkage system.*  
Ms. **Marcela Quintero**, Data Analyst, Telethon Institute for Child Health Research, Centre for Child Health Research, the University of Western Australia; Western Australian Department of Health (Perth).

13.00–13.45 Lunch break

### **Second session: Data linkage outcomes for research and policy**

(Chairperson: Mr. Michael Ruland, NEPS, WZB)

13.45–14.55 *Social inequalities in preterm births and associated economic cost.*  
Dr. **Amanda Langridge** and Mr. **Jeff Cannon**, Telethon Institute for Child Health Research, Centre for Child Health Research, the University of Western Australia (Perth).

15.00–15.50 *Entering and leaving home care during childhood: A Cumulative incidence study of groups most at risk.*  
Dr. **Melissa O'Donnell** and Ms. **Miriam Maclean** (PhD candidate), Telethon Institute for Child Health Research, Centre for Child Health Research, the University of Western Australia (Perth).

15.50–16.15 Coffee break

16.15–16.55 *Using mapping to identify socio-demographic inequalities in poor mental health and early child development.*  
Ms. **Janice Wong** (PhD candidate) and Mr. **Scott Sims**, Telethon Institute for Child Health Research, Centre for Child Health Research, the University of Western Australia (Perth).

17.00–17.30 *Socio-economic disadvantage, parental mental health and deliberate self-harm and suicidal death in young people: a birth cohort study using record linkage.*  
Mr. **Nan Hu** (PhD candidate), Telethon Institute for Child Health Research, Centre for Child Health Research, the University of Western Australia (Perth).

## **Wednesday, 4 September 2013**

**Data collection systems in Germany for health and social science research: current situation and future trends**  
(Chairperson: Dr. Roland Habich)

9.00–9.30 *Actors, tasks and structure in the field of nationwide empirical research in Germany.*  
Dr. **Roland Habich**, Head of Data Management, WZB (Berlin).

9.30–10.00 *The German Data Forum (RatSWD) – improving the research data infrastructure in Germany.*  
Mr. **Denis Huschka**, M.A., Managing Director, RatSWD (Berlin).

- 10.00-10.30 *Thirty years of “living in Germany” – The history (and future development) of the German Socio-Economic Panel (SOEP).*  
Prof. Dr. **Jürgen Schupp**, Director, SOEP (Berlin).
- 10.30-11.00 Coffee break
- 11.00-11.30 *Monitoring social inequality and health – The Federal Health Reporting Service (GBE).*  
Dr. **Lars Kroll**, Researcher, Robert Koch Institute, GBE (Berlin).
- 11.30-12.00 *Education as a lifelong process – The German National Educational Panel Study (NEPS).*  
Mr. **Daniel Bela**, NEPS, University of Bamberg.
- 12.00-13.00 Lunch break
- 13.00-13.30 *SHARE-RV: Linking survey data of SHARE with social security data of the German Pension Insurance.*  
Ms. **Christin Czaplicki**, M.Sc., Record Linkage Project, Munich Centre for the Economics of Ageing (MEA), Max Planck Institute for Social Law and Social Policy.  
Ms. **Tatjana Mika**, Head of the Research Data Centre, German Pension Insurance (RV).
- 13.30-14.00 *Linkage of health data with administrative labour data of the Institute for Employment Research (IAB).*  
Mr. **Stefan Bender**, Head of the Research Data Centre, Federal Employment Agency, IAB (Nuremberg).

**WITH SUPPORT FROM:**



## **Australian data linkage infrastructure: building on the West Australian experience**

*Dr Merran Smith*

Australia has a range of good quality, population-based information resources on health and human services. While these data collections are generally held in organisationally separate information systems, they can be linked using privacy preserving methodology. Systematic population-based data linkage began in Western Australia (WA, current population of 2.5 million people) in 1995. After a number of challenges including access to funding and concerns about privacy, the WA data linkage system has developed into a valuable state asset covering four decades of linked data on births, deaths, electoral roll, early childhood development, education, community development and justice. The linked data has supported a wide range of important population health research that has changed health policy, improved clinical practice and delivered better patient outcomes. New South Wales (population around 7.5 million) implemented a similar system in 2006. In 2009, the Australian Government funded the Population Health Research Network (PHRN) to coordinate the development of a data linkage infrastructure across Australia. Australia now has the capacity to link data within and between data collections held by the Australian government and state and territory governments. In this presentation I share with the audience our challenges and successes in developing Australian data linkage, an innovation in research and policy.

### **Dr. Merran Smith**

BSc, MSc, BA, PhD, Post Grad Dip Management. Dr. Smith was appointed as the inaugural Chief Executive of Australia's Population Health Research Network in April 2009. The Network has received \$A 33 million over six years from the Australian Government through the National Collaborative Research Infrastructure Strategy, the Super Science Initiative and the Collaborative Research Infrastructure Scheme to build data linkage capability within and between Australian jurisdictions. Jurisdictions, agencies and academic partners are contributing a further \$41 million in cash and in-kind to development of the capability. Prior to taking up this appointment Dr Smith worked in the Western Australian government sector where she was a member of the Senior Executive Service for more than 15 years. She was Director of the Western Australian Department of Health's Health Information Centre for more than a decade. The Centre housed Western Australia's major state-wide health data collections including the hospital morbidity, perinatal, mental health and cancer registry collections. She was responsible for establishing data linkage as a core Department of Health service during this period. Dr Smith has served as Chair or Member of a number of Australia's peak national health information committees including the Statistical Information Management Committee and the National Public Health Information Working Group. She continues to actively contribute to information policy and strategy at state and national level, and is committed to the on-going development and use of Australia's health information and related resources for system improvement.

## **Linking the Western Australian Health Data System to Social Outcome Data: Why, how, and what came out of it?**

*Dr. Jianghong Li*

Western Australia (WA) has developed a unique health data linkage system that connects information about all health and social events for individuals, including about 20 million hospital admission records (dating back to 1970) which are linked to birth, marriage, electoral records, cancer, mental health, and death. In 2004 this enormous health linkage system was augmented to be linked with social outcome data on education, disability, juvenile delinquency and child protection. Since 2004 this cross-government jurisdictional data linkage has evolved into what is called today the "Developmental Pathways in WA Children Project (DPP)", encompassing even broader linkages with data on housing, drug use, and police arrests, involving 11 state government agencies as research partners. In this presentation, I provide a detailed account on the scientific and policy rationale behind the linkage expansion beginning in 2004 (the genesis of the DPP), research and policy outcomes emanating from the cross-jurisdictional data linkage, and the funding mechanisms that facilitated the linkage.

**Dr. Jianghong Li** completed her Bachelor Degree of Arts with a major in English and English Language from Yunnan University in 1982 in China. She obtained her Master's Degree in 1986 and PhD in Sociology 1992 in the United States. She was awarded the Mellon Postdoctoral Fellowship in 1992-1993 at the University of Washington in Seattle. Between 1994 and 1996 she was a Research Associate at the Swiss Federal Institute of Technology in Zurich, conducting research on further education and occupational mobility in Switzerland. Between 2000 and 2012 she conducted research in social epidemiology in Perth, Western Australia, as a Senior Research Fellow at Curtin University and an Associate Professor at Telethon Institute for Child Health Research, the University of Western Australia.

Dr. Li was the Guest Editor-in-Chief for a special issue on Social Determinants of Child Health and Wellbeing published in *Health Sociology Review* June 2009. She was an Associated Editor of *Rural Sociology*, published by the American Rural Sociological Society (June 2005-2010). She has supervised four PhD students, one honours student and three Master students to completion. She is currently supervising three PhD students, and is a Senior Researcher at Wissenschaftszentrum Berlin für Sozialforschung.

Dr. Li's broad research interests include social, economic, psychological, cultural influences on child health and development; the interface between socio-economic, psychological and biological factors contributing to inequalities in child health and development. She has published widely on the health and development of Australian children and has also published research on gender inequality and health and drug use/HIV in adolescents in China. She has extensive experience in collaborating with scientists from psychology, nutritional epidemiology, labour economics and biomedical sciences. She is interested in the broad issue of modernity's paradox whereby we witness declines in indicators of human development and rising social inequality despite postmodern society's unprecedented economic prosperity and technological advancement. Her current research focuses on the connection between the 24 hours/7 day economy and the health and development of children in Australia and other developed countries. Specifically, she is interested in the impact of parental employment status, commuting to work, work hours and work scheduling (e.g., non-standard work schedules) and job characteristics on children's health and development (e.g., behavioural problems, diet and obesity, risk taking, and educational achievement) and factors that may mediate or moderate this impact, such as family resources and processes.

### **Technical and practical aspects of cross agency data linkage: the West Australian example**

*Prof. Nicholas de Klerk*

Encouraged by results from the Oxford Record Linkage Study, linking of administratively collected health data started in Western Australia in the 1970s, but in piecemeal fashion, funded by various research grants. The system was more firmly established in the 1990s and has progressed and expanded since then. The success of the system and its ability to expand into linking data from many government agencies outside of health, has been based on several established principles: use of the data for research purposes only; the separation of the linkage information, linkage processes and linkage personnel from data collation and delivery; and engagement of all stakeholders (including consumers and the general community) in both setting the research agenda and also maintaining a strong governance structure. Maintaining these strong partnerships ensures a continuation of the mutual trust between stakeholders and has enabled excellent research and excellent communication of research results and their practical implications.

**Prof. Nicholas de Klerk** leads the Biostatistics and Genetic Epidemiology division at the Telethon Institute for Child Health Research. He is Winthrop Research Professor at the Centre for Child Health Research, the University of Western Australia, and Adjunct Professor in the School of Population Health, with over 30 years' experience in biostatistics and epidemiology, and numerous publications and successful research grants in record linkage and using linked datasets.

## **Linking data to build an evidence base: Developmental Pathways in WA Children Project**

*Dr. Rebecca Glauert*

The Developmental Pathways in WA Children Project is a landmark project taking a multidisciplinary approach to investigate the pathways to health and wellbeing, education and juvenile delinquency outcomes among Western Australian children and youth. To achieve this, researchers from the Telethon Institute for Child Health Research and the University of Western Australia have been working in collaboration with 13 state government departments, including the WA Departments of Health, Education, Child Protection, Corrective Services, Communities, Indigenous Affairs, Treasury and Finance, Housing, Attorney General, Training and Workforce Development, Disability Services Commission, Mental Health Commission, and WA Police. The project has established the process of linking together de-identified, longitudinal, population-based data collected and stored by a large number of the WA government departments and the Telethon Institute, to create a world class research and policy planning/evaluation resource. The hope is that through the effective communication of the research findings, future government agency policies, practice and planning initiatives will be more preventative, culturally appropriate and cost efficient. This presentation will provide an outline of the Project; demonstrate best practice for large, cross government linkage projects; discuss the importance of using administrative data in policy making; and provide an update on some of our research findings.

**Dr. Rebecca Glauert** graduated from the University of Western Australia, School of Psychology with a PhD in 2008. Following the completion of her PhD, she was employed as a Program Manager to manage the Developmental Pathways in WA Children Project (DPP), the largest cross jurisdictional data linkage project in Australia. The DPP links together information from the WA departments of Health, Education, Child Protection, Corrective Services, Disability Services, with others such as Housing, Police, and Attorney General coming on board to link their data in the coming years. Through the use of linked cross jurisdictional administrative data, the project aims to understand the trends and factors that influence both positive and adverse life outcomes for children, including poor mental health, disability, child maltreatment, school failure, and juvenile delinquency. Dr. Glauert's role includes the advocacy of both data linkage, and the importance of using cross agency government data to inform, evaluate and improve policy making, practice, prevention and intervention. Additionally, she works in both a researcher role, as well as a liaison between researchers and policy makers, ensuring that research findings are translated into policy relevant briefings. The project is an excellent example of the benefits of having researchers, policy makers and practitioners working together with the common goal of improving child health and wellbeing.

Dr. Glauert's research interests focus around child development and Mental Health. She supervises 3 PhD students in the areas of deliberate self-harm and suicide, education, and mental health. Her work focuses on understanding the individual, family and community risk factors for educational outcomes, deliberate self-harm, suicide and poor mental health, from birth through adolescence. Dr. Glauert is a very strong believer in the benefits of cross-sectoral research, and the importance of knowledge transfer, and believes that the only way to effect change in our society, and to improve the worrying outcomes for children and youth, is to have a universal approach, one that joins up policy makers, researchers, heads of government, practitioners, social service sectors, and community members. Dr. Glauert believes that effective information and knowledge transfer are key to any research having an impact on policy, practice and prevention.

## **Developmental Pathways Project: Challenges for the WA Data Linkage system**

*Ms. Marcela Quintero*

Data linkage research is a fundamental part of societal development. The changing nature and complexity of research projects consistently observed over the last decade have introduced significant challenges to data linkage models. Balancing communication, confidentiality, privacy and timely data release have become a core issue for researchers, data custodians and the community. This presentation will illustrate some of the complexities of the Developmental Pathways Project (DPP) and will present some of their effects on the Western Australian (WA) data linkage model.

The very unique nature of the research undertaken by DPP has introduced several challenges to the data linkage model in WA. Firstly, the project involves a large collaboration of government agencies and sources data from all of them. Consequently, it has required the creation of a linkage infrastructure which has been maintained since the inception of the project, nine years ago. This infrastructure involves linkage of numerous databases otherwise not accessible for research. The increased number of datasets utilised in a single project and a wider use of population-based cohort designs has significantly broadened the complexity of the research. Such complexity has been associated with added constraints for approving and releasing data due to privacy and confidentiality issues. Furthermore, the process of data extraction prior to its release usually implies duplicate efforts from the data collections involved. The combination of the above issues has led to less efficient processes for delivery of data and the timely commencement of studies.

Joint efforts among agencies have been made to deal with the growing complexity of the project and to counteract the implications of more rigorous approval processes. Some of the strategies carried out include the establishment of new communication conduits and development and implementation of data resources for custodians and researchers. These resources include a tool for centralising service data extractions and cohort selection (CARES) and a metadata website containing information about the data sources. In addition to this, a consumer reference group has been gradually set up to involve community members. Forthcoming additions to the already established challenges comprise the expansion of the infrastructure by adding datasets from Police and the Department of Housing. It is also expected that negotiations will commence with other non-health government agencies to include their data into CARES.

The above strategies aim to deal more efficiently with the issues raised by the growing demands of complex research projects. They have also contributed to gradually raising awareness among researchers, custodians and consumers of the impact the evolving nature of data linkage research has on balancing confidentiality, privacy and the release of data. It is expected the new strategies sought and implemented in the near future will contribute to addressing the demands of emerging research intricacies.

**Ms. Marcela Quintero** holds a joint position between the Western Australian Department of Health and the Telethon Institute for Child Health Research. She received her Medical degree from the Universidad Nacional de Colombia and completed a MPH at Curtin University of Technology in Western Australia. Her current research interest focuses mainly on data linkage and the area of juvenile offending.

### **Social inequalities in preterm births and associated economic cost**

*Dr. Amanda Langridge and Mr. Jeffery Cannon*

Preterm birth is associated with a range of childhood morbidities and in industrialised societies is the primary cause of infant mortality. Social and racial inequalities in preterm birth have been reported in North America, UK, Europe and New Zealand. This study utilised population-level data to investigate social and racial inequalities in preterm birth, as well as cost of preterm birth, among Aboriginal and non-Aboriginal infants in Western Australia. All live, singleton births between 1984 and 2010 were included, and multivariable logistic regression was used to investigate relative differences in preterm infants between socio-economic groups. Aboriginal and non-Aboriginal infants were analysed separately. Inequalities in preterm births between Aboriginal and non-Aboriginal infants increased over time, with the percentage of preterm births being almost two-fold higher for Aboriginal infants (14.8%), compared with non-Aboriginal infants (7.6%). A significant portion of the disparity between Aboriginal and non-Aboriginal infants is attributable to parental socio-economic and demographic characteristics, though the disparity continues to persist even after adjustment for these factors. While the overall rates of preterm birth in Western Australia have remained fairly static over the last two decades, the disparity between Aboriginal and non-Aboriginal infants has increased and is now similar to inequalities seen 20 years ago. These findings highlight a major public health issue that should be of great concern, given the short- and long-term morbidities and complications associated with preterm birth.

**Dr. Amanda Langridge** is a social epidemiologist, with a particular interest in perinatal and child health, and she was awarded her PhD, entitled "Social and racial inequalities in birth rates and infant outcomes in Western Australia" in December 2009. Her thesis was examined by a leading social epidemiologist from Harvard University, who predicted that upon publication of her papers on monitoring inequalities in preterm birth and poor fetal growth, they would become seminal papers for the future monitoring of equity of birth outcomes. Dr Langridge's current areas of research include preterm birth, intellectual disability and autism spectrum disorders.

**Mr. Jeffery Cannon** currently works as a health economist and data analyst at the Telethon Institute of Child Health Research in Perth, Western Australia. His primary interest is in the use of decision analytic modelling techniques to investigate the impact of health care interventions. These models are used to inform stakeholders on the cost-effectiveness of prospective changes to health care policies.

### **Entering and leaving home care during childhood: A cumulative incidence study of groups most at risk**

*Dr. Melissa O'Donnell and Ms. Miriam Maclean*

In Australia, like many developed countries, increasing numbers of children are being placed in out-of-home care due to child maltreatment. Concerns have been raised about the pressure on out-of-home care systems to keep up with increasing numbers of children requiring care. Given these increases, our study assesses what proportion of children experience out-of-home care during their childhoods, the characteristics of families with children taken into care, and how demographic and socio-economic indicators predict the likelihood of a child entering care. This study utilises linked longitudinal population level data from the Western Australian Departments of Health and Child Protection.

**Dr. Melissa O'Donnell** is an NHMRC Early Career Fellow and a Psychologist, who completed her PhD in 2009 and is based at the Telethon Institute for Child Health Research and the University of Western Australia. Her research uses linked longitudinal population data to investigate emergency department presentations and hospital admissions related to child abuse and neglect; determine outcomes of children who have contact with the child protection system; and investigate the child, family and community characteristics which increase or reduce vulnerability to child abuse and neglect.

**Ms. Miriam Maclean** is a PhD student and Research Assistant on the Developmental Pathways project, focusing on the developmental outcomes of maltreated children including those in out-of-home care. Prior to joining the Telethon Institute for Child Health Research, she completed a Bachelor of Arts and Masters of Science in psychology, and worked for 7 years as a researcher in the public and private sectors.

### **Using mapping to identify socio-demographic inequalities in poor mental health and early child development**

*Ms. Janice Wong and Mr. Scott Sims*

The environment in which a child develops has been identified to play a large role in shaping the trajectory of their health, education and social outcomes. This study focusses on mental health outcomes, and outcomes measured by the Australian Early Development Index (AEDI); an index that measures five key areas of early childhood development: physical health and wellbeing, social competence, emotional maturity, school based language and cognitive skills, and communication skills and general knowledge. Using logistic regression, we identified the significant risk factors associated with poor mental health and children who are developmentally vulnerable (i.e., low AEDI scores). These risk factors, typically socio-demographic in nature, were then geographically mapped to the metropolitan area of Perth, Western Australia. The mapping of these outcomes and risk factors can help identify the areas of highest need and priority for intervention.



**Ms. Janice Wong** is completing her doctorate on the Developmental Pathways Project (DPP). Her project is titled "The relationship between educational and mental health outcomes for Western Australian children: A longitudinal population study". Using linked longitudinal population data provided through the DPP, this subproject seeks to explore the dynamic relationship between children's educational outcomes and their mental health, whilst taking into account variables that have been shown to impact on this relationship.

**Mr. Scott Sims** is a Data Analyst on the Development Pathways Project at the Telethon Institute of Child Health Research. He recently completed a Master of Biostatistics at University of Sydney and has also developed an interest in spatial analysis using linked cross-jurisdictional data.

### **Socio-economic disadvantage, parental mental health and deliberate self-harm and suicidal death in young people: a birth cohort study using record linkage**

*Mr. Nan Hu*

**Aims:** This study aims to investigate whether or not and how restricted foetal growth, familial and neighbourhood socio-economic disadvantages in early life, parental psychiatric and DSH presentations may interact to influence the risk of DSH related hospital presentations among adolescents, and how these effects would be modified by child gender.

**Methods:** A nested case-control sample was compiled from a birth cohort using administrative health data collected by the Western Australian (WA) government. Up to 2011, 2,142 people with DSH related hospital presentations were identified among people born alive during 1991-1999 in WA, and 42,840 controls were matched using incidence density sampling method by gender, birth year, and date of the first DSH presentation of children.

**Results:** Socio-economic disadvantages in early life (living in more disadvantaged neighbourhoods, born to teenage parents, born to a single parent family, and high parity) and parental psychiatric and DSH presentations independently increased offspring's risk of DSH related hospital presentation in adolescence. These factors also accounted for some of the effects of low birth weight and preterm or over mature birth on the outcome. The effect of parental presentations prior to the birth of the child and during the infancy was modified by child gender. Generally, mother's psychiatric and DSH presentation had a greater effect on children than that of father's, and girls were more vulnerable than boys to parental psychiatric and DSH presentation.

**Conclusions:** The findings provide insights into the mechanisms of how social and biological determinants interact to shape the risk of more severe DSH behaviour of adolescents that requires more medical attentions.

**Mr. Nan Hu** is a PhD student of the University of Western Australia, working on the Developmental Pathways Project to undertake his research on deliberate self-harm (DSH) behaviour among children and young adolescents using administrative linked data. He has a Master's Degree in Statistics and 5-year working experience in public health.

### **The German Data Forum (RatSWD) - improving the research data infrastructure in Germany**

*Mr. Denis Huschka, M.A.*

In the digital age, there are more potentially interesting data available than ever before. However, the more data becomes available, the harder it is to find relevant data for research purposes which are essential for innovative and robust research. Finding these data should be as easy as finding information on Google. The presentation elaborates on why it is beneficial to share data, how sharing should be organized and which challenges lie ahead. Sharing data facilitates essential replication studies, enables researchers to make the most of limited resources, and gives their research more exposure and thus a higher reputation. Especially in the social sciences, the process of sharing data requires an established infrastructure that guarantees easy access to individual data for researchers while safeguarding data protection and high quality of data. In Germany, this is done by a growing, specialized infrastructure made up of research data centres under the auspices of the German Data Forum. Although much has been accomplished there are many challenges still ahead.

**Mr. Denis Huschka** is Managing Director of the German Data Forum (RatSWD). He studied Sociology and Political Science at the Berlin University of Technology, where he earned his Master's Degree. Mr. Huschka is also Permanent Visiting Fellow of the German Socio-Economic Panel Study (SOEP) / German Institute for Economic Research (DIW Berlin) and Research Associate of the Institute of Social and Economic Research (ISER), Rhodes University Grahamstown, South Africa.

### **Thirty years of "living in Germany" – The history (and future development) of the German Socio-Economic Panel (SOEP)**

**Prof. Dr. Jürgen Schupp** studied Economics and Sociology in Mainz and Frankfurt am Main, completing his diploma in Sociology at the Johann Wolfgang Goethe University in Frankfurt in 1983. He then took over a position as Junior Research Associate at the University of Frankfurt in the German Research Foundation's Special Research Unit "Microanalytic Foundations of Social Policy" (Sfb 3). At the end of 1984, he transferred within the Special Research Unit into the project "Socio-Economic Panel", located at Deutsches Institut für Wirtschaftsforschung, DIW (German Institute for Economic Research), under then-president Hans-Jürgen Krupp, and thereafter made a series of longer research visits to the USA and Great Britain. In 1994, he completed his doctorate at the Ruhr-University Bochum. He has worked closely with the Free University Berlin for many years as a visiting lecturer at the Institute for Sociology. He joined the Forschungsinstitut zur Zukunft der Arbeit, IZA (Institute for the Study of Labour), as a Research Fellow in October 2000. In March 2006 Dr. Schupp was appointed Honorary Professor for Sociology in the Faculty of Political and Social Sciences at the Free University and in 2008 as Vice-Dean of Graduates Studies at the DIW Graduate Center of Economic and Social Research. From 12/2005-03/2006 and 08-09/2006 he worked as Research Fellow at the Hanse Institute for Advanced Study (HWK) in Delmenhorst. As survey manager and deputy director of the longitudinal socio-economic panel study (SOEP), Jürgen Schupp has been responsible for planning the further development of SOEP as a survey instrument for many years. In February he was appointed Head of Research Unit Socio-Economic Panel Study (SOEP).

### **Monitoring social inequality and health – The Federal Health Reporting Service (GBE)**

*Dr. Lars Kroll*

The Monitoring of Social and Health Inequality at the Robert Koch Institute (RKI) combines the results of the Federal Health Reporting in Germany with those of the Federal Poverty Monitoring to analyse the extent and the development of health inequalities in Germany. The Federal Health Monitoring System was established in 2008 and consists of health examination surveys and health interview surveys with longitudinal components being alternately performed for adults as well as for children and youth (DEGS, KiGGS). These surveys are continuously accompanied by annual cross-sectional interview surveys of the population living in Germany (GEDA). The Health Surveys contain a considerable number of socio-demographic characteristics operationalized with the concepts of poverty reporting in mind. For the Federal Health Reporting results of the Health Surveys of the RKI and of various other national and international sources are combined to form a comprehensive picture of health status of the German population as well as the health status of social groups (i.e. population at risk of poverty, unemployed, lone mothers, etc.).

**Dr. Lars Kroll** holds a BA in Sociology from Freie Universität and a PhD in Social Sciences from Humboldt Universität Berlin. He has worked since March 2009 as Researcher in the Department 2 Epidemiology and the Federal Health Reporting Service at the Robert Koch Institute in Berlin. His research interests focus on social inequality and health, social change, regional disparities in health, and work and health.

## **Education as a lifelong process - The German National Educational Panel Study (NEPS).**

*Mr. Daniel Bela*

The National Educational Panel Study (NEPS) develops and conducts panel surveys in six age cohorts. The data produced by these panel studies are, accompanied by two additional cross-sectional studies on school reforms, edited, documented and disseminated to the scientific community. As a result, NEPS Scientific Use Files establish a comprehensive base for research on educational trajectories, their outcomes and implications. Researchers from various disciplines (including sociology, psychology, educational research, economics and demographics) are already using NEPS data to further insights into educational development. In order to further enhance the usefulness of the data, NEPS, in close cooperation with the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB), works on linking the surveys' data to administrative employment data held at the FDZ. This includes efforts in probabilistic record linkage to identify a respondent's records in the administrative data. As a result, the NEPS Data Center and the FDZ will jointly provide linked data in the future.

**Mr. Daniel Bela**, born in 1983, studied social sciences at the University of Erlangen-Nuremberg and holds a corresponding diploma (Dipl.-Sozialw. Univ.). His research interests focus on applied informational sciences in social research. Since 2010, he has been working as Research Assistant at NEPS Data Center, which is part of the working unit data and methods within NEPS. As a data manager, he is entrusted with the NEPS' Scientific Use File data edition, dissemination and documentation.

## **SHARE-RV: Linking survey data of SHARE with social security data of the German Pension Insurance**

*Ms. Christin Czapliski and Ms. Tatjana Mika*

In order to improve data quality and to reduce respondents' burden, survey data are increasingly often linked to administrative data. SHARE-RV is a data linkage project that stands for the direct linkage of survey data from the Survey of Health, Ageing and Retirement in Europe (SHARE) with administrative records of the German Pension Insurance (Deutsche Rentenversicherung) using the Social Security Number (SSN) as unique identifier. The SSN is an identifier used by the German Pension Insurance and the Employment Agency. The goal of SHARE-RV is to link survey with administrative data to provide a rich database for researchers from different fields. Linkage requires the respondents' written consent within the SHARE interview. Given that the survey is computer assisted (CAPI), an additional paper form was necessary to collect all needed information and the signature. After collecting, documenting and checking this information, the linkage is carried out.

So far the linkage has been performed for wave 3 and wave 4 of SHARE. The survey's third wave, SHARELIFE, collected detailed retrospective life histories of Europeans 50 years old and over. The combination of accurate administrative data and profound information about different aspects of the respondents' lives provides a wide range of research possibilities. Since SHARELIFE covers a long episode of life, beginning in the early childhood until present, the temporal overlap between SHARELIFE and the administrative records is quite high. Additional to the enrichment of the SHARELIFE data with very detailed longitudinal information on lifelong earnings, for example, the fact that some information is included in both data sets (like the job history) offers the possibility to validate the retrospectively collected SHARELIFE data.

The presentation compares information on work history and income in order to show the benefit of having multiple data sources. For this purpose we use retrospective information and compare this with records of the German Pension Insurance of the same person. As a result we are able to provide deeper insights in the accurateness of retrospectively recorded survey data, as well as how survey data enrich administrative records.

**Ms. Christin Czaplicki** studied at the University of Rostock (Germany) and holds a BA in Social Sciences, specialized in empirical social science, statistics, demography, sociology and adult education; and a Master's Degree in Demography, specialized in fertility, mortality, migration and statistical data analysis. Since 2010 she has been working as a researcher in the Project "SHARE-RV: Analysis of old-age provision, health and family" at the Munich Centre for the Economics of Ageing in the Max-Planck-Institute for Social Law and Social Policy. Ms. Czaplicki is also since 2011 a lecturer at Technische Universität Berlin, Department for Social Sciences, division for empirical research.

**Ms. Tatjana Mika** is since 2012 Head of the Research Data Centre of the German Pension Fund, where she started working as a researcher back in 2004, and since 2007 she works also as Lecturer at the Humboldt University and Technische Universität Berlin, Department for Social Sciences, division for empirical research. Ms. Mika holds a Master's Degree in Sociology, and has also studied Law.

### **Linkage of health data with administrative labour data of the Institute for Employment Research (IAB)**

*Mr. Stefan Bender*

The goals of the German Record Linkage Center (German RLC) are to promote research on record linkage techniques and to increase the number and quality of record linkage applications in empirical research in Germany. Although the German RLC is currently funded by the German Research Foundation, the center is designed as a long-term infrastructure facility. Due to the broad applicability of record linkage, the center is open for all academic disciplines. The two goals of the German RLC are pursued in cooperation between the University of Duisburg-Essen (Institute of Sociology) and the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB).

The German RLC provides the following services: consulting of linkage projects, conducting data linkages, providing a clearing house for linkage of sensitive data and maintaining a website as an information portal for potential users. Furthermore, the center disseminates its record linkage Software Merge ToolBox (MTB), which is free for academic purposes.

Besides these services, the German RLC conducts various research projects on the following central topics of record linkage methodology: error-tolerant record linkage techniques with encrypted identifiers (Privacy Preserving Record Linkage), blocking with encrypted identifiers using multi-bit-trees and the development of a new error-tolerant and anonymous linking code (Cryptographic Longterm Key).

While the nexus of health and epidemiological research with employment research has only started to gain importance in Germany, the IAB, the FDZ and the German RLC already collaborate with two studies in this field: one is the project "Occupational health research in epidemiological cohort studies" (AeKo), a part of the on-going "Heinz Nixdorf Recall Study" (HNR). The project's aims are to link its participants' survey responses and examination results with their occupational biographies included in the administrative data of the IAB and to validate the occupational biographies contained in both data sources against each other. In a further step, AeKo aims at deriving occupational risk factors from these longitudinal data.

The second study is "lidA-leben in der Arbeit" (German Cohort Study on Work, Age and Health), a longitudinal study that examines the effects of work on older employees' health. This study has also been linked to the administrative employment data of the IAB as well as to health insurance data.

Based on the German RLC's expertise in the field of data linkage, the IAB and the FDZ are currently consulting the planned National Cohort, a large-scale, nationwide, long-term population study, on their plans to conduct data linkage with a variety of administrative data sources within the strict borders of German data protection legislation.

**Mr. Stefan Bender** holds a BA in Sociology and a MA in Sociology/Statistics from the University of Mannheim (Germany). He is since 2005 Head of the Research Data Centre of the Institute of Employment Research, Federal Employment Agency, where he worked as research assistant between 1992 and 2005. He was also research assistant in the Department of Sociology at the University of Mannheim between 1990 and 1992.