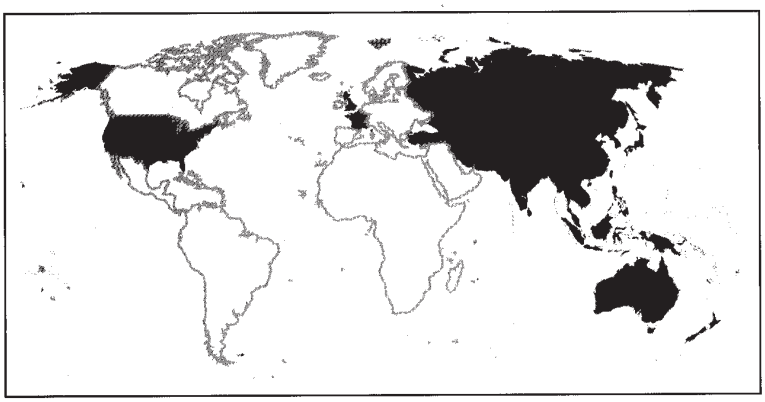
Asia-Pacific Population Journal

Vol. 23, No. 3 ISSN 0259-238X December 2008

United Nations
ESCAP

ECONOMIC AND SOCIAL COMMISSION FOR ASIA AND THE PACIFIC

ESCAP is the regional development arm of the United Nations and serves as the main economic and social development centre for the United Nations in Asia and the Pacific. Its mandate is to foster cooperation between its 53 members and 9 associate members. ESCAP provides the strategic link between global and country-level programmes and issues. It supports Governments of the region in consolidating regional positions and advocates regional approaches to meeting the region's unique socio-economic challenges in a globalizing world. The ESCAP office is located in Bangkok, Thailand. Please visit our web site at <www.unescap.org> for further information.



The shaded areas of the map indicate ESCAP members and associate members.

COVER PHOTOGRAPH

A Hmong grandmother and her granddaughter in a village in the Lao People's Democratic Republic (Photograph by Susan Hardman).

The passage of time, sometimes over many decades and generations, is precisely at the core of this special issue of the *Asia-Pacific Population Journal*, dedicated entirely to the topic "Understanding Health and Population over Time".

Full text of articles available at:
http://www.unescap.org/appj.asp

ASIA-PACIFIC POPULATION JOURNAL Vol. 23, No. 3, December 2008

CONTENTS

The Asia-Pacific Population Journal is published three times a year in English by the United Nations Economic and Social Commission for Asia and the Pacific (ESCAP).

Produced by the Social Development Division of ESCAP, APPJ provides a medium for the international exchange of knowledge, experience, ideas, technical information and data on all aspects of the field of population, in order to help developing countries in the region improve the utilization of data and information for policy and programme purposes, among others.

The designations employed and the presentation of the material in this periodical do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Mention of firm names and commercial products does not imply the endorsement of the United Nations.

ESCAP is not accountable for the accuracy or authenticity of any quotations from sources identified in this publication. Questions concerning quoted materials should be addressed to the sources directly.

ST/ESCAP/2523

Foreword

Over the years, longitudinal studies in Asia and the Pacific have made a unique contribution to the understanding of health, demographic and social trends and the causal factors contributing to them. The knowledge and insight that such studies reveal are necessary to measure large-scale changes and their impact on policy, thus helping societies to better address the long-lasting and the emerging challenges confronting them. However, a persistent lack of capacity and resources, especially in resource-poor settings, has largely constrained the undertaking of such studies and analysis, as well as the utilization of the data generated.

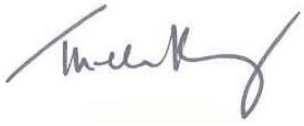
In order to highlight the scientific and policy relevance of longitudinal studies and advocate greater investment in this type of research, the Institute for Population and Social Research (IPSR), Mahidol University, Thailand, organized the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific in Bangkok on 24 and 25 May 2007. The event was a part of the Institute's broader efforts to promote research into the health and demographic changes in the region that had been initiated in 2000. The conference, co-sponsored by the Welcome Trust under its Health Consequences of Population Change Programme and the Rockefeller Foundation, provided a forum to discuss the state of the art of longitudinal studies in Asia and the Pacific. The two-day meeting brought together a wide range of experts from various countries and institutions in the region and beyond. Papers presented during the conference assessed the role of longitudinal studies in monitoring and analysing population and health trends (including epidemiological concerns, ageing and lifestyle patterns), and the actual utilization of these studies. The papers also reflected on the specific contribution made by such studies to science and their potential for scientific as well as policy purposes. The need to develop strategies to strengthen human and technical capacity in longitudinal data collection and analysis was also highlighted.

This special issue of the Asia-Pacific Population Journal *on understanding health and population over time is aimed at conveying a sense of the discussions held at the conference to a wider public of interested key readers, planners, decision makers, policymakers and academics. The seven articles included herein present an overview of longitudinal studies in Asia and the Pacific, and one in comparison to similarly constrained settings such as in parts of Africa, providing evidence of their theoretical*

and applied relevance. The challenges posed by their complex, labour-intensive and expensive nature are discussed and examples of cohort models and ongoing research efforts in Asia are provided. The contribution made by those research studies to action-oriented policies is also highlighted; a more technical aspect of longitudinal studies is also tackled, namely that of database management systems.

The selection of articles published here, from more than 50 presentations made during the conference, and the peer review of the various drafts were entrusted to an editorial team led by Rosalia Sciortino and composed of Philip Guest, Sureeporn Punpuing and Amara Soonthorndhada – all associated at the time with IPSR. We are greatly indebted to them. Without their diligent efforts and the active cooperation of the authors throughout the peer-review and editing processes, the publication of this special issue would not have been possible.

As essential research tools, longitudinal studies are key instruments for improving the understanding of ongoing population and health transformations and crafting more effective policies. It is our hope that this extrabudgetary issue of the Journal*, focused upon a largely overlooked methodological issue, will help establish the potential of longitudinal studies for the advancement of health and population sciences and for improving the welfare of people in the region.*



Thelma Kay

Director
Social Development Division
Economic and Social Commission
for Asia and the Pacific (ESCAP)

Rosalia Sciortino

Chair of the Conference
Institute for Population
and Social Research
Mahidol University (IPSR)

Special Issue

Understanding Health and Population over Time

Abstract

Longitudinal studies in the Asian and Pacific region show a wide diversity of research designs and modes of utilization. In drawing on these experiences, the article discusses the relative advantages of different approaches and the multiple ways in which they have contributed to building demographic and health knowledge and to shaping related policies, interventions and services, thus contributing to the welfare of people in the region. Because of this theoretical and applied relevance, the authors conclude that the undertaking of longitudinal research should be encouraged, notwithstanding the challenges posed by their complex, labour-intensive and expensive nature, and the extra precautions needed to address ethical concerns.

The history of longitudinal community studies in Africa attests to their value for demographic and health research, policy and practice and their resilience in the face of multiple challenges. This collection provides a comparative parameter for

activities in the Asian and Pacific region. The article discusses the contributions of longitudinal community studies to demographic surveillance and health research in Africa and considers the many challenges faced when undertaking them in resource-poor settings. In Africa, where these studies are most needed, the shortage of technical, human and financial resources has required sites to collaborate, share technology and pursue common strategies, giving impetus to methodological and technical innovations.

The Cebu Longitudinal Health and Nutrition Survey (CLHNS) is the longest health and nutrition longitudinal study ever conducted in the Philippines and possibly one of the longest in Asia. It started in 1983 with 3,327 pregnant women who, together with their children born between May 1983 and April 1984, have been reinterviewed 18 times thereafter. This article discusses the origins and modalities of CLHNS, underscoring its value in examining social, demographic and health trends. The challenges faced and the lessons learned during the survey's more than two decades of existence are presented. Best practices are described and a set of recommendations for the conduct of longitudinal and intergenerational studies is also put forward.

Since its establishment in 1985, the Primary Health Care Network of (the Islamic Republic of) Iran has provided a unique opportunity for annual longitudinal data collection on the total population of a large number of rural communities. The village-based rural health workers who run the system are given necessary training to collect a variety of data on the population of the main and satellite villages served by them. The main instrument of data management used is a sheet of paper 50 x 70 cm with concentrically coloured circles at the centre; it is known as the "vital horoscope". In addition to graphically representing monthly events, the vital horoscope includes seven tables summarizing the results of the annual census, detailed information about live births, neonatal, infant, child and maternal mortality, under-five mortality by cause of death, and family planning. This article presents the main findings of an analysis of vital horoscopes covering

16,000 rural health houses representing all provinces of the Islamic Republic of Iran during the period 1993-2003. The data indicate impressive changes in the age structure, fertility rates and various health indicators of villagers covered by the primary health-care system.

The Effects of Intergenerational Support on the Changes in Cognitive Functioning of the Rural Elderly in China **75**

The migration of working-age adults from rural areas of China has altered traditional patterns of intergenerational support among elderly persons who remain in the rural areas. This study examines how intergenerational support has influenced the changes in cognitive functioning of the rural elderly in that country. Data derived from a 3-wave longitudinal survey from 2001 to 2006 of 997 parents aged 60 and older living in rural areas of Anhui Province are used in the study. Hierarchical linear models are used to estimate the effect of intergenerational transfers of financial, instrumental and emotional support on the development of the cognitive functions of the elderly parents. After controlling for the needs of the elderly, two factors, namely receiving financial support from adult children and providing financial support to adult children, contributed to changes in the cognitive functioning of the elderly; moreover, there are some direct relationships between the other time-varying supports and the changes in cognitive functioning of the elderly. These results suggest that intergenerational transfers, especially financial transfers, have a significant influence on changes in the cognitive functioning of the elderly in rural China. In the context of the high levels of migration of working-age adults, the results have clear implications concerning the effects of public policy on the social support system.

Evaluation of the Thai Tobacco Control Policy **91**

Tobacco control has become an important health policy for the Government of Thailand. Among the policies introduced in Thailand in the past three years are increased taxes on tobacco, graphic pictorial labelling on cigarette packets and bans on the advertising of cigarettes. However, there is little international evidence to suggest that these policies are effective. For this reason, Thailand has joined an international project that collects longitudinal data to evaluate tobacco-control policies. This study presents a comparison of data from the first two rounds of the project in Thailand. Although there are several promising signs of success in the

attempts by the Government of Thailand to reduce the use of tobacco, this study also indicates some concerns. The proportion of persons smoking hand-rolled cigarettes has increased. The article suggests that this is probably a result of tax increases on manufactured cigarettes. Most of the sample agreed with the prohibitions on smoking in hospitals, offices, restaurants and other public places. The placing of graphic warning labels on cigarette packets has met with success. Knowledge of the health impacts of smoking increased and most smokers reported that the graphic warning labels were most effective in making them think about the health dangers of smoking.

Managing the Kanchanaburi Demographic Surveillance System: Creation of a Relational Database Management System 107

The Kanchanaburi Demographic Surveillance System (KDSS) monitors population change in Thailand's Kanchanaburi Province. KDSS was established with a cross-sectional scheme for data collection and manipulation. This original database needed to be transformed into a longitudinal scheme in order to achieve productive data utilization as the original KDSS data structure was a "flat file" in SPSS format and was difficult to manage and manipulate. Three main problems with the original database were identified. First, the linkage between the various data sets required excessive human resources. Second, some data were missing as a result of individuals moving from one study area to another during consecutive years. Last, but not least, a number of data retrievals were inconsistent because the data capture was employed separately on the survey of the previous year. The new system is operated using a relational database management system, which is compatible with the spatial data. The procedures of the transformation diminish data redundancy, unreliability and inconsistency. Hence, the human resources needed for data retrieval are reduced, data quality is improved and data confidentiality is strengthened.

Editorial Management

Editor-in-Chief: Keiko Osaki

Editor: Wanphen Sreshthaputra

Advisors: Bhakta Gubhaju, K.S. Seetharam, Jerrold W. Huguet and Ghazy Mujahid

Guest editors: Rosalia Sciortino, Philip Guest, Sureeporn Punpuing and Amara Soonthorndhada

Editorial Advisory Board

BENCHA YODDUMNERN-ATTIG, Associate Professor, Institute for Population and Social Research, Mahidol University, Nakhon Pathom, Thailand

BHASSORN LIMANONDA, Director, College of Population Studies, Chulalongkorn University, Bangkok

SUREEPORN PUNPUING, Director, Institute for Population and Social Research, Mahidol University, Nakhon Pathom, Thailand

G. GIRIDHAR, Special Advisor, Asia-Pacific Regional Office (APRO) and UNFPA Representative in Thailand

GRACE CRUZ, Director, Population Institute, University of the Philippines, Quezon City, Philippines

JOHN KNODEL, Professor, Population Studies Center, University of Michigan, Ann Arbor, United States of America

NIBHON DEBAVALYA, former Division Director, United Nations Economic and Social Commission for Asia and the Pacific, Bangkok

SHIREEN JEJEEBHOY, Senior Programme Associate, Population Council, New Delhi

VIPAN PRACHUABMOH, Principal Investigator, Asian MetaCentre, Singapore

WASIM ZAMAN, Executive Director, International Council on Management of Population Programmes (ICOMP), Malaysia

Longitudinal Research Designs and Utility in the Asian and Pacific Region

The experience of the Asian and Pacific region shows a great diversity in longitudinal research designs, the findings of which are benefiting the understanding of health and demographic transformations and contributing to more effective policies, programmes and services.

By Bencha Yoddumnern-Attig, Philip Guest, Varachai Thongthai, Sureeporn Punpuing, Chanya Sethaput, Aree Jampaklay, Rossarin Gray and Yupin Vorasiriamorn*

Longitudinal research, which includes panel research, is the term used to differentiate the methodology and utility of prospective studies from that of cross-sectional research. It describes not a single method, but a family of methods that measure change by linking individual data across time (Zazzo, 1967, cited in Menard, 2002).

* All authors were associated with the Institute for Population and Social Research, Mahidol University, Salaya, Nakhon Pathom, Thailand, at the time of first submission of the article to the *Asia-Pacific Population Journal*. E-mail: prbyd@mucc.mahidol.ac.th.

Longitudinal research is valuable for understanding the causes and effects of socio-economic changes stemming from globalization; and population trends associated with changes in morbidity, mortality and migration; as well as health and epidemiological transitions within countries, especially at the micro level, such as at the individual, family or community levels. In the field of health, for example, one prominent concern is how to measure a population's health accurately, especially when it is facing many dynamic social, economic and behavioural changes that affect health (Murray and others, 2002). For their part, policymakers need to understand the dynamics that underlie key development issues in order to address them properly at policy and service levels (Rose, 2000).

Longitudinal research designs have thus become essential research tools for understanding change. As Ruspini (2002) noted, since the data are collected about the same population at different points of time, longitudinal research is able to present information about what happened to a set of units (people, households, firms etc.) across time. In this way, longitudinal data can also help to clarify ambiguities in causal relations through the temporal ordering of events and experiences in people's lives (Scott, 1995).

In this article, the authors provide an overview of longitudinal research design and utility in the Asian and Pacific region which is partly based on the articles presented at the International Conference on Understanding Health and Population over Time, which was organized in Bangkok on 24 and 25 May 2007 by the Institute for Population and Social Research.

An overview of longitudinal research designs

Longitudinal research designs often combine both extensive (quantitative) and intensive (qualitative) approaches. They can also allow for cross-sectional or longitudinal data analysis (Kalton and Citro, 2000) and qualitative multiple-case study analysis (Yoddumnern-Attig and others, 2006).

In the Asian and Pacific region, longitudinal population surveillance systems were established in the late 1960s, the most prominent being the Ballabgarh site in Haryana, India, in 1965 (Kapoor and others, 2006) and the Matlab site in the Chandpur District of Bangladesh in 1966 (Razzaque, 2007). Matlab surveillance system data, in particular, have enabled researchers to gain a much fuller understanding of the dynamics of family planning and child survival (Arends-Kuenning, 2002; Bairagi, 2001; Phillips and Hossain, 2003). Several other surveillance systems in the Asian and Pacific region were established thereafter, and some of their designs can be classified under four major types: (a)

repeated cross-sectional survey or trend analysis; (b) panel design; (c) longitudinal community studies; and (d) record linkages or administrative panels.

Before describing them in more detail below, it is important to stress that, irrespective of their variety in focus, covering the individual, a group of individuals (cohort), households and communities, as well as record linkages, *it is individuals who are followed over time*. In household panels, for example, data about household dynamics come from individuals who are associated with the households and their families (Rose, 2000; Ruspini, 2002). Another shared feature is that, given the length of the research and the need for follow-up, all longitudinal studies ought to apply strict ethical safeguards and ensure confidentiality to protect the respondents. It is also extremely important that respondents understand the implications of the longitudinal design and that they be contacted repeatedly at regular intervals.

In relative terms, the appropriateness of the research design is judged and adopted based on the particular research problem (Menard, 2002), since each design has specific purposes, merits and disadvantages, as will become clear from the description of the previously identified four major types presented below. Awareness of their relative strengths and weaknesses is important in order to develop appropriate strategies for reducing problems and maintaining the benefits of the design.

Repeated cross-sectional survey or trend analysis

In a repeated cross-sectional survey design, *the same question is regularly asked over a period of time to a different sample population at each time interval.* This design is extremely common and is most favoured by descriptive studies that focus on identifying changing trends. Much of our knowledge of trends in demographic and health outcomes derive from repeated cross-sectional surveys. For example, Thailand's contraceptive prevalence surveys (CPSs) adopted this design. The data from five survey rounds were compared in order to measure changes in contraceptive use and the fertility of women aged 15-49 during the period 1978-1996. The results indicated a continued decline in the total fertility rate, although at a slower pace in later years (Chamratrithirong and others, 1997; Leoprapai and Thongthai, 1989).

This design's major strength lies in its ability to measure change at an aggregate level. Changes in survey estimates and trends reflect changing population characteristics and changes in population composition (Kalton, 2004). However, this design is much less useful when it comes to determining causal relations because events cannot be precisely ordered over time.

Panel design

In a panel design, *the same individuals or groups of individuals are repeatedly interviewed at regular intervals* often using the same questions. Although this general principle is practised, panel designs can vary in the way the panel is established, in the means to maintain sample characteristics and in the ways to reduce distortion through the loss of subjects owing to factors such as mortality, boredom, mobility and refusals. Major modes of panel design and their features include the following types of panel.

Rotating panels are usually established through probability sampling. *Individual panel members are rotated in and out of the panel where they are followed at a specified interval for a relatively short period of time. New members are added to the new sample at each successive wave*. Such rotation enables the original characteristics of the sample to be maintained and reduces the distortion caused by the loss of subjects. In addition, a major feature of the rotating panel design is its ability to combine the features of both panel and cross-sectional analysis (Ezzati-Rice and Cohen, 2004; Rose, 2000). In the United States of America, Ezzati-Rice and Cohen (2004) utilized a rotating panel to develop predictive models for expenditures on medical care, yielding important information on the persistence of high medical expenditure costs, the precursors of health outcomes and possible cause-effect relationships as correlated with rapidly changing health-care scenarios.

Split panels are established *through long-term panel members who are followed over time as well as through the use of an additional replenishment sample.* A replenishment sample is added at each successive wave to ensure adequate sample size and survey continuity (Menard, 2002). Because it is similar to the rotating panel, the split panel design can be combined with a cross-sectional study design (Kalton and Citro, 2000). Sirirassamee and others (2007) used this design in their longitudinal data on Thailand's Tobacco Control Policy (2005-2009) to measure gross changes in responses and to evaluate the effectiveness of the policy, including the use of graphic warning labels on packets of tobacco in particular. Study results indicate that, while some success is evident in reducing the use of tobacco, an increase in the use of hand-rolled cigarettes has also been noted, possibly due to their lower price and absence of a warning label. This shift in tobacco consumption suggests the need for strengthened campaigns to influence smokers of factory-made and hand-rolled cigarettes so that they will quit smoking or reduce their intake of tobacco.

Cohort panels are designed *to observe intensively people who share a common life event at two or more times*. Commonly used cohorts include birth cohorts, people who enrolled in a programme in the same year, people who are in the same age

cohort and people who retired in a one- or two-year time frame (Neuman, 2006). The main focus of the cohort panel design is on examining the cohort and its important features, not on specific individuals. Its strength lies in its ability to identify the temporal ordering of sample members' experiences and to analyse the relationships of earlier experiences with later outcomes (Kalton, 2004); for example, it can highlight the interaction between early life experience and adult outcomes, while its weaknesses are most notably in terms of its longevity and the fixed sample structure (Wadsworth and others, 2003).

For the cohort panel design, replacement does not normally take place, which can lead to a decline in sample size over time. To minimize sample loss and distortion, several strategies can be used, such as birth cohort, childhood, catch-up and middle-life strategies, while tracking strategies are fundamental to all (Wadsworth and others, 2003). The *birth cohort strategy,* as illustrated by Choprapawan (2007) in her study of Thai children and families, and by Feranil, Gultiano and Adais (2007) in their discussion of the Cebu Longitudinal Health and Nutrition Survey in the Philippines, entails collecting data starting at birth and continuing thereafter into adulthood. In the *childhood strategy*, as used by Avila (2007) to measure the impact of the absence of mothers on the development of Filipino adolescents, data collection starts in childhood and continues into adult life. The *catch-up strategy* entails collecting data in childhood and then again in adulthood, as noted by Dempster-McCain and Moen (1998). In a *middle-life strategy,* as adopted by Byles and Dobson (2007) in their presentation of the findings from the Australian Longitudinal Study on Women's Health, and Seubsman and others (2007) in their national cohort study of the Thai health-risk transition, follow-up begins with cohort members who are in middle life and retrospective data are collected about their earlier experiences.

Household panels are based on initial probability sampling of individuals or households drawn from a population in a given area and following this population over time. Repeated interviews are carried out at a fixed interval and pay particular attention to the behaviours of individuals, families and households over time (Ruspini, 2002). Often all adult members in the household (not only a household head) are also repeatedly interviewed in successive waves. *The aim is to discover what happens to the same subject over a certain period of time, thus making it possible to study microsocial change.* When individuals are studied over time, it is possible to explore the dynamics of individual and family behaviours as they are affected by economic, demographic and social changes as well as the strategies they use to respond to such changes (Anderson, Bechhofer and Gershuny, 1994; Ruspini, 2002; Scott, 1995). Often-cited studies from outside the Asian and Pacific

region are, for example, the British Household Panel Study, the German Socio-economic Panel and the United States Panel Study of Income Dynamics.

Longitudinal community studies

While the individual or household is the main focus of other panel studies, the longitudinal community study *concentrates on a community as a whole (a total population), covering every household and individual living in a well-defined location* (Mbacke and Phillips, 2007; Mehryar, Naghavi and Kasemipour, 2007; Razzaque, 2007). Complete registration of all households is essential to explore, prioritize and address community issues in a timely and equitable manner (Taylor, 1997). Repeated data collection is carried out at fixed intervals, depending upon the project. A community panel is established by following individuals and households in the same community across time.

The demographic surveillance system (DSS), or the continuous recording of all births, deaths and migrations (Baiden, Hodgson and Binka, 2006), is a major component of several longitudinal community studies. Design variations exist, however, owing to differences in fundamental objectives as well as approaches to population health.

DSS sites, such as the two Vietnamese field laboratories of Chililab in Chi Linh District of Hai Duong Province and Filabavi in Bavi District of Hatay Province, the Matlab site in Chandpur District of Bangladesh and the Purworejo site in Central Java, Indonesia, use a public health approach to develop intensive small-scale vital registration systems and to serve as a comprehensive source of demographic and health information for the populations under study. Such information can be used to identify solutions to major health and population problems. Continuous programme monitoring and evaluation can be accomplished through such a rich source of longitudinal evidence (Anh and others, 2006; Long and others, 2006; Mehryar, Naghavi and Kasemipour, 2007; Razzaque and Streatfield, 2006; Wilopo and others, 2006).

Thailand's Kanchanaburi Demographic Surveillance System (KDSS), in Kanchanaburi Province on the border with Myanmar, has instead used a social approach to understand the relationship between socio-economic and demographic changes and the health consequences of the population in the study area (IPSR, 2004). KDSS is a hybrid design. While adopting the main features of longitudinal community studies, it incorporates distinctive features of household panel studies. Although data are collected annually for every household, all adult members in a household are repeatedly interviewed in successive waves (Guest and Punpuing, 2003). By linking data at individual, household and community levels over time

(across survey rounds), changes at each of these levels can be identified, such as changes in economic or health conditions, changes in household and community structures and processes, as well as changes in individual outcomes.

No matter what approach a project uses, *the design of longitudinal community studies is aimed at exploring demographic and health events in the same community over a certain period of time.*

Record linkages or administrative panels

The most important feature of this design is the *linkage between longitudinal survey and administrative cross-sectional sequences.* Data items that are not collected primarily for panel purposes are linked with administrative records using specific personal identifiers. The value added by this particular design comes from the pooling of data from several sources, which enables researchers to resolve methodological shortcomings in an efficient manner (Ruspini, 2002; Trivellato, 1999).

An ongoing project on malaria prevention among the Karen[1] population in the upland area of Kanchanaburi Province is a good example of this type of design (Yoddumnern-Attig and others, 2006). The KDSS data gathered annually from 2000 to 2004 are linked with data from the malaria registration record administered by the Malaria Division, Department of Communicable Disease Control, Ministry of Public Health. Data analysis can produce incidence and prevalence rates that vary annually, as well as provide an understanding of possible cause-effect relationships that stem from social, economic and behavioural changes over time.

Utility of longitudinal studies in the Asian and Pacific region

Longitudinal designs are used widely in population and health research, especially for studies that require a proper understanding of microlevel changes. They can also serve as valuable tools for initiating, or strengthening, a comprehensive vital registration system, and as a source of vital demographic and health information for populations living in particular priority areas (Mbacke and Phillips, 2007; Razzaque and Streatfield, 2006; WHO, 2006). In addition to these broad frames, longitudinal studies have also been used for several other specific purposes.

Longitudinal studies can generate accurate cohort data for estimating changes in basic demographic and health parameters, such as birth rates, fertility rates, maternal mortality rates and infant mortality rates). An analysis of longitudinal data from Ballabgarh in India, Matlab in Bangladesh and the Primary Health Care Network of the Islamic Republic of Iran, for example, indicate

comparable patterns of change in basic demographic and health parameters, such as neonatal, infant and under-five mortality rates (Kapoor and others, 2006; Mahryar, Naghavi and Kasemipour, 2007; Razzaque and Streatfield, 2006).

Longitudinal data can be used to develop a tool to obtain accurate information for a valid measure of mortality. As an integral part of KDSS, Prasartkul and others (2007) developed a verbal autopsy tool and computer software program under the patented name of "The Mahidol Verbal Autopsy System", which enables users to obtain valid information on the causes of death in different settings.

Longitudinal panel samples can be used to ascertain the health trends of a population. Byles and Dobson (2007) have used longitudinal data on women's health in Australia to measure health trends. Overweight, with adverse long-term health consequences, has been identified as a major health problem among Australian women as they age. Similar evidence has also been found among the elderly in Viet Nam and the recommendations made for public health interventions (Nguyen, 2007).

Longitudinal data have been used to develop predictive models for providing elderly persons with support. Studies from Bangladesh, China, the Philippines and Thailand, for example, have focused on developing behavioural models to measure health transitions, cognitive functioning, living arrangements and care-giving patterns among the elderly. Evidence from China and Thailand indicate that changes in the living arrangements of the elderly, as triggered by child migration, might disrupt traditional types of care-giving for the elderly (Min and Punpuing, 2007; Ping and Shu-Zhou, 2007). When social and financial supports are examined, several studies showed that children's migration can have a positive impact on the health of the elderly (Kuhn, 2003; Ping and Shu-Zhou, 2007; Punpuing and Guest, 2006). Cruz and Agustin (2007) highlight an important issue based on their study of Filipino elderly persons, that is, a realistic health transition model needs to take into account movement into and out of one's initial health status.

Longitudinal data can be used to examine relationships between variables across time. In analysing KDSS data on parental absence owing to migration and children's school enrolment, Jampaklay (2006) noted that adverse effects arise when the parents move. In particular, her results suggested that the long-term absence of mothers appears to reduce the educational chances of children left behind, and that maternal roles are not easily replaced by those of other family members.

Longitudinal research can measure gross flows into and out of a particular state. For example, So (2007) used the Cambodian panel data set to analyse household dynamics and movements into and out of poverty. Study findings provide valuable information on poverty assessment and the recognition that poverty reduction will depend on the success of broadening economic development into rural areas and securing rights of access to common property resources for the poor.

Longitudinal panel samples can assess health intervention effectiveness and living standards. Oliveras, Nahar and Johnston (2007) used the Matlab database from 1982 to 2005 to explore the impact of changes in Bangladesh's contraceptive delivery strategy. In Indonesia, Pitoyo (2007), using longitudinal data from the Indonesian Family Life Survey, assessed the impact of the Asian financial crisis (1997-2000) on the living standards of a sample population. Results showed a mixed picture in terms of people's standards of living. Moreover, the argument that the crisis affected macroeconomic performance more than household and individual economics seems to be upheld by the study's findings.

Longitudinal data can be used to develop appropriate feedback models for community utilization. For example, Gray, Pleumcharoen and Suwannopakao (2007) used KDSS data to develop communication models through which the data could be practically and effectively used at community and beneficiary levels.

Conclusion

Well-designed longitudinal studies have been used in many ways in the Asian and Pacific region. The studies presented above reflect the growing need for, and the prevalence of, longitudinal research designs in the region. Worthy of particular note is the proliferation of longitudinal community research designs. Found across the region, these designs have devoted particular attention to documenting relationships in the field of population and reproductive health. Overall, longitudinal studies are providing the valuable data necessary to establish causal relationships, thus providing a much better understanding of how rapid social changes affect the health of individuals and populations. Because of their theoretical and applied relevance, the undertaking of longitudinal research designs should be encouraged, notwithstanding the challenges posed by their complex, labour-intensive and expensive nature, and the extra precautions needed to address ethical concerns. While there is still much to be done to meet these challenges, the wide range of longitudinal research being conducted in the region is benefiting the understanding of health outcomes as an essential step in developing effective population and health policies, programmes and services in many countries of the region.

Endnote

1. An ethnic minority group originally from Myanmar currently living in areas of Thailand close to the Myanmar border.

References

Anderson, M., F. Bechhofer and J. Gershuny (eds.) (1994). *The Social and Political Economy of the Household,* Oxford, United Kingdom, Oxford University Press.

Anh, L.B. and others (2006). "Chililab DESS", *INDEPTH Monograph*, vol. 1, part C.

Arends-Kuenning, M. (2002). "Reconsidering the doorstep-delivery system in the Bangladesh family planning program", *Studies in Family Planning,* vol. 33, No. 1, pp. 87-102.

Avila, J. (2007). "Mother absence and development of Filipino adolescents", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Baiden, F., A. Hodgson and F.N. Binka (2006). "Demographic surveillance sites and emerging challenges in international health", *Bulletin of the World Health Organization,* vol. 84, No. 3, pp. 161-256.

Bairagi, R. (2001). "Effects of sex preference on contraceptive use, abortion and fertility in Matlab, Bangladesh", *International Family Planning Perspectives,* vol. 27, No. 3, pp. 137-143.

Byles, J. and A. Dobson (2007). "Findings from the Australian longitudinal study on women's health", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Chamratrithirong, A., P. Prasartkul, V. Thongthai and P. Guest (1997). *National Contraceptive Prevalence Survey 1996,* Institute for Population and Social Research, Mahidol University, publication No. 212.

Choprapawan, C. (2007). "The longitudinal cohort study of Thai children and family: Pitfall and benefit for low and middle income countries", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Cruz, G.T. and S. Agustin (2007). "Determinants of health transition among Filipino older people", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Dempster-McCain, D. and P. Moen (1998). "Finding respondents in a follow-up study", in J.Z. Giele and G.H. Elder, Jr., eds., *Methods of Life Course Research: Qualitative and Quantitative Approaches,* Thousand Oaks, California, United States, Sage Publications.

Ezzati-Rice, T. and S.B. Cohen (2004). "Design and estimation strategies in the medical expenditure panel survey for investigation of trends in health care expenditures", in S.B. Cohen and J.M. Lepkowski, eds., *Eighth Conference on Health Survey Research Methods,* Hyattsville, Maryland, United States, National Center for Health Statistics.

Feranil, A.B., S.A. Gultiano and L.S. Adair (2007). "The Cebu Longitudinal Health and Nutrition Survey: Two Decades After", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Gray, R., S. Pleumcharoen and R. Suwannopakao (2007). *Research and Development to Develop Appropriate Models for Information Utilization,* Institute for Population and Social Research, Mahidol University, publication Nos. 334 and 335 (in Thai).

Guest, P. and S. Punpuing (eds.) (2003). *Report of Kanchanaburi Demographic Surveillance System Round 2* (2001), Institute for Population and Social Research, Mahidol University, publication No. 275, Salaya, Nakhon Pathom, Thailand.

IPSR (2004). *Strengthening Institutional Capacity in Population and Health Research and Training in the Asian Region,* report of a regional seminar, 10-12 November, Bangkok, Institute for Population and Social Research, Mahidol University, publication No. 292.

Jampaklay, A. (2006). "Parental absence and children's school enrollment: Evidence from a longitudinal study in Kanchanaburi, Thailand", *Asian Population Studies,* vol. 2, No.1, pp. 93-110.

Kalton, G. (2004). "Introduction to session 1: Capturing diversity and change in a dynamic population", in S.B. Cohen and J.M. Lepkowski, eds., *Eighth Conference on Health Survey Research Methods,* Hyattsville, Maryland, United States, National Center for Health Statistics.

_____and C.F. Citro (2000). "Panel surveys: Adding the fourth dimension", in D. Rose, ed., *Researching Social and Economic Change,* London and New York, Routledge.

Kapoor, S.K., K. Anand, B. Nongkynrih and E. Paul (2006). "Ballabgarh DSS", *INDEPTH Monograph,* vol. 1, part C.

Kuhn, R. (2003). "A longitudinal analysis of health and mortality in a migrant-sending region of Bangladesh", paper presented at the International Workshop on Migration and Health in Asia, Singapore.

Leoprapai, B. and V. Thongthai (1989). "Fertility and family planning in Thailand", *Journal of Population and Social Studies,* vol. 2, No. 1, pp. 21-41 (in Thai).

Long, N.H. and others (2006). "Filabavi DSS", *INDEPTH Monograph,* vol. 1, part C.

Mbacke, C. and J.F. Phillips (2007). "Longitudinal community studies in Africa: Challenges and contributions to health research", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Mehryar, A., M. Naghavi and S. Kasimipour (2007). "Vital horoscope: An innovative method for longitudinal data collection on demographic characteristics and health status of rural population developed by Iranian Primary Health Care System", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Menard, S. (2002). *Longitudinal Research,* London, Sage Publications.

Min, Q. and S. Punpuing (2007). "Labor migration and elderly living arrangement change in Kanchanaburi Demographic Surveillance System (KDSS), Thailand", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Murray, C.J.L., J.A. Solomon, C.D. Mathers and A.D. Lopez (2002). "Summary measures of population health: Conclusions and recommendations", in *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications,* Geneva, World Health Organization.

Neuman, W.L. (2006). *Social Research Methods. Qualitative and Quantitative Approaches*, sixth edition, Boston, Massachusetts, United States, Pearson International Edition.

Nguyen, M. (2007). "Trends of nutritional status, its associated factors among ageing Vietnamese", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Oliveras, E., L. Nahar and H.D. Johnston (2007). "The impact of contraceptive service delivery strategies on pregnancy termination in Bangladesh", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Phillips, J. and M. Hossain (2003). "The impact of household delivery of family planning services on women's status in Bangladesh", *International Family Planning Perspectives,* vol. 29, No. 3, pp. 138-146.

Ping, W. and L. Shu-Zhuo (2007). "The effects of intergenerational supports on the change of cognitive function of rural elderly in China", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Pitoyo, A.J. (2007). "Indonesian health standards: The evidence of dynamic health conditions from the Indonesian Family Life Survey", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Prasartkul, P., Y. Porapakham, P.Vapattanawong and J. Rittirong (2007). "Development of a verbal autopsy tool for investigating cause of death", *Journal of Population and Social Studies,* vol. 15, No. 2, pp. 1-22.

Punpuing, S. and P. Guest (2006). "Migration and agricultural production in Kanchanaburi, Thailand", paper presented at the INDEPTH Second MUWG Analytic Workshop, Johannesburg, South Africa, 28 August - 1 September.

Razzaque, A. (2007). "Different methods of data collection for population and health research in developing countries: Strength and challenges", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

_____ and P.K. Streatfield (2006). "Matlab DSS", *INDEPTH Monograph,* vol. 1, part C.

Rose, D. (ed.) (2000). "Household panel studies: An overview", in *Researching Social and Economic Change: The Uses of Household Panel Studies,* London and New York, Routledge.

Ruspini, E. (2002). *Introduction to Longitudinal Research,* London and New York, Routledge.

Scott, J. (1995). "Using household panels to study micro-social change innovation", *The European Journal of Social Sciences,* vol. 8, issue 1, March.

Seubsman, S., A. Sleigh and the Thai Cohort Team (2007). "The Thai health-risk transition: A national cohort study", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Sirirassamee, B. and others (2007). "Evaluation of the Thai tobacco control policy: Results from wave 1 and wave 2 surveys", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

So, Sovannarith (2007). "Moving out of poverty with particular reference to nine rural villages", paper presented at the International Conference on Understanding Health and Population over Time: Strengthening Capacity in Longitudinal Data Collection and Analysis in Asia and the Pacific, Bangkok, 24-25 May.

Taylor, C.E. (1997). "Origins of longitudinal community-based studies", in M. Das Gupta and others, eds., *Prospective Community Studies in Developing Countries,* Oxford, United Kingdom, Clarendon Press.

Trivellato, U. (1999). "Issues in the design and analysis of panel studies: A cursory review", and "Longitudinal analysis: A bridge between quantitative and qualitative social research", *Quality and Quantity,* vol. 33, No. 3, special issue, July-August.

Wadsworth, M.E.J. and others (2003). "The life course perspective design: An example of benefits and problems associated with study longevity", *Social Science and Medicine,* vol. 57, No. 2003, pp. 2193-2005.

WHO (2006). Household and facility surveys (online), Available from: www.who.int/healthmetrics/documents/hmnissue_householdfacilitysurveys.pdf, accessed on 26 May 1007.

Wilopo, S.A. and others (2006). "Purworejo DSS", *INDEPTH Monograph,* vol. 1, part C.

Yoddumnern-Attig, B., C. Sethaput, K. Rojkureestien and S. Pleumcharoen (2006). "Multimethod research for malaria prevention and control in upland area, Kanchanaburi Province", paper presented at Participatory Prevention and Control : PPC Organized by the Center of Disease Control, Region 8, Ministry of Public Health at Pimarn Hotel, Nakhon Sawan, Thailand, 3-4 August (in Thai).

Longitudinal Community Studies in Africa: Challenges and Contributions to Health Research

The history of longitudinal community studies in Africa attests to their value for demographic and health research, policy and practice, and their resilience in the face of multiple challenges. Their potential to contribute significantly to monitoring progress towards attaining the Millennium Development Goals should be harnessed.

By Cheikh S.M. Mbacké and James F. Phillips*

Reliable and timely information is a central pillar of well-functioning health systems. In fact, much of the progress in improving human health is "due to technical progress, including advances in knowledge about diseases and about appropriate, cost-effective responses" (Jamison and others, 2006; p. 155). Developing health technology requires sound research and the precise information that is needed for gauging the efficacy of interventions. Moreover, comprehensive sociodemographic and health information is needed to guide policy deliberations. The need for such information is particularly acute in

* Cheikh S.M. Mbacké, independent consultant, e-mail: csmbacke@orange.sn, and James F. Phillips, Heilbrunn Department of Population and Family Health, Mailman School of Public Health, Columbia University, New York, United States of America.

developing countries because policy that lacks empirical grounding can be more damaging to public welfare in such settings than is the case in relatively prosperous countries where the public enjoys multiple service options and individuals have resources for financing personal choices.

In most developed countries, the vital registration and health data that are needed to guide health policy and action are comprehensive, available and complete. However, in sub-Saharan Africa, the coverage of the vital registration and health systems is inadequate. Household surveys are the most important source of information on population health, but retrospective data fail to provide essential information on the timing of events. The causes and consequences of problems remain poorly understood; thus, policies must be formulated in an information vacuum.

Longitudinal data-collection efforts attempt to reproduce the wealth of information that is afforded by the continuous recording of events and their causes. Prospective analyses of longitudinal data permit scientists to examine the incidence of events, change over time and the causes of change, all of which are fundamental requirements of scientific inference that are very difficult with cross-sectional studies.

The critical need for longitudinal investigation justifies the cost and complexity of tracking individuals over time. Longitudinal health research falls into three general categories, all of which involve "following" a precisely defined population over an extended period of time: (a) cohort studies which follow a subgroup of individuals selected on the basis of a time-specific characteristic such as year of birth; (b) panel studies which follow a representative sample of households whose members are tracked and reinterviewed over time; and (c) longitudinal community studies which cover every household and individual living in a well-defined location.

Unlike panel and cohort studies, which are interested in subgroups of the population, longitudinal community studies focus on the community as a whole, even if the data are collected on an individual basis. Longitudinal community studies have at their core a demographic surveillance system (DSS), which records continuously and meticulously all births, deaths and migrations, the only sources of change in the community's initial population.[1]

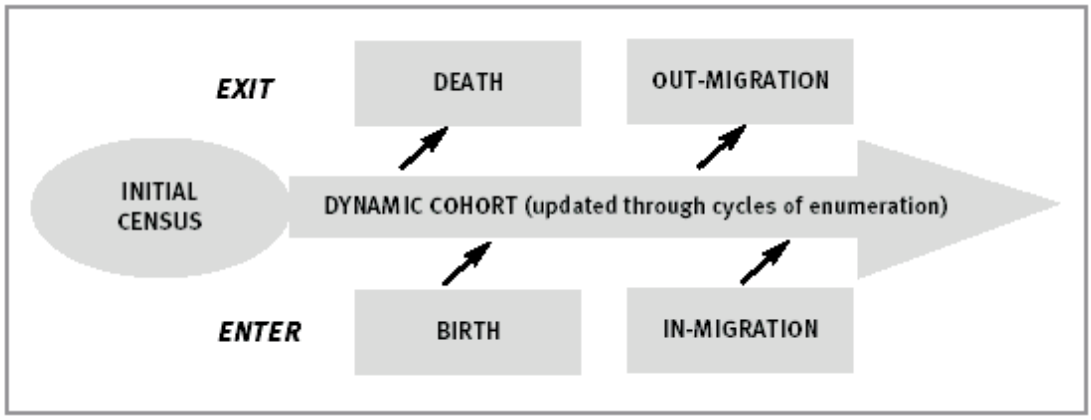
The purpose of this article is to discuss the contributions of longitudinal community studies to health research in Africa as well the challenges faced by the scientists and decision makers who rely on them to guide health policy and practice. The authors begin with an overview of the DSS concept and turn to a

discussion of the history of its use in sub-Saharan Africa, the contribution of this legacy to the science of demographic surveillance and the technical challenges that are motivating DSS methodological initiatives in Africa. Particular attention is devoted to the emergence and function of the INDEPTH Network, an international network of DSS sites involved in demographic and health research; the network is based in Accra.

State of the art of the demographic surveillance system

DSS starts with a baseline census that identifies each and every household and individual residing in a well-demarcated area. A system is then put in place to record continuously all births, deaths and migrations (of individuals and households) that occur in this area and their exact timing. The product that emerges is an "open cohort", a longitudinal register of individuals at risk of events as demographic dynamics unfold over time (see figure 1). The typical DSS is augmented with the registration of marriages, divorces and changes in status and household relationships, additional information that is critical to achieving an understanding of the demographic dynamics that are being observed.

Figure 1. The "Open Cohort" concept



Depending on the objectives of a study, this core system is enriched by additional information collected from the same population. The most advanced studies add verbal autopsies of all deaths in order to identify underlying causes of death and to track the burden of disease.

Advanced DSS systems define the open cohort information in figure 1 at the individual level. This capability depends upon assigning a unique identification number to every individual and household that enables tracking of all elements of change so that data from all studies in the DSS area can be logically related to the individuals under observation. The fact that each study can build on the foundation

provided by previous studies permits the accumulation of scientific capabilities and serendipitous research. Well-designed cross-sectional surveys, for example, link to the figure 1 open cohort data, permitting studies that draw upon survey information as covariates of the prospective events that transpire in the time following cross-sectional data collection. Some of the more remarkable contributions of longitudinal community studies were completely unanticipated by the designers of cross-sectional studies that have been fielded in DSS sites.

In order to realize the full potential of DSS, exposure to risk must be well defined for each individual. This requires information on the timing of migration and procedures for linking out-migration records to individuals at risk and procedures for initiating the observation of in-migrants at the time of their arrival in a DSS household. This requirement complicates data management significantly but it also enables research that fulfils the requirements of demanding impact studies involving hazard models and individual-level inference. Since health research aimed at determining the effectiveness of a given drug or intervention tolerates little ambiguity in exposure to the risk of getting the target disease, there is no shortcut to recording unambiguously all entries (birth, in-migration) and exits (death, out-migration) and their timing in order to ensure the reliability of estimates of exposure, incidence and lethality. Naïve attempts to simplify data collection by dropping any of the demographic events in figure 1 lead to incomplete information on the population at risk and impaired capability to conduct individual-level analysis. By addressing this need, well-designed DSS-based population and health studies permit causal inference and formal appraisal of hypotheses on the impact of interventions. Initiating and sustaining a good DSS is challenging, however, requiring interdisciplinary cooperation of medical, computer and population scientists.

Demographic surveillance system in Africa

The history of DSS in sub-Saharan Africa attests to both the inherent importance and challenges associated with the longitudinal approach to data collection. Although the underlying concepts of surveillance were first demonstrated in China and Europe, Africa has recently been at the forefront of innovation in scientific demographic surveillance.

The idea and practice of collecting registration of vital events in defined populations was well developed during China's Ching dynasty in the sixteenth century, and systematic longitudinal registration dates back 2,000 years to China's Han dynasty, although these early efforts were limited to the coverage of adults. For the most part, archival historic records were compiled for administrative or religious functions, such as the European parish registers of the seventeenth to

nineteenth centuries (Phillips, 2003). Modern longitudinal community studies for research investigations emerged from a synthesis of public health, epidemiology and demography in the early twentieth century. The first such study on record is the pellagra study (1916-1921) designed to identify the etiology of this nutritional disorder in a number of poor and deprived villages in South Carolina, United States of America (Das Gupta and others, 1997). In referring to such studies, Omran (1974) concluded that the application of epidemiological methods and techniques to the field of population dynamics and family planning created a new field of science that he referred to as "population epidemiology".

Methods of surveillance that combine data management, linkage and scientific inference were first developed in Africa as products of studies focusing on the epidemiology of tropical diseases. The first such study was launched by the Pholela Health Centre in 1940 to study the diseases that were prevalent in rural Natal (South Africa) and workable methods for their prevention and treatment (Ngom and others, 2001). The Pholela study is credited with having collected longitudinal DSS-type data for a population of 10,000 over a period of 15 years. Although it was closed in 1955, its methods and productivity as a research station inspired much of the innovation in community longitudinal research that followed (Ngom and others, 2001).

In 1950 another Gambian epidemiological study investigating the effects of disease in a tropical environment was launched by Sir Ian McGregor of the British Medical Research Council in the Gambia (McGregor and Smith, 1952). The Keneba study, as it is known, set up a system whereby all births and deaths occurring in three villages (Keneba, Kanton Kunda and Manduar) were continuously recorded by literate village residents. This project added the innovation of monitoring non-vital events and critically important characteristics of the study population, such as detailed information on marriage, migration and other residential changes, and anthropometric and health status for all residents.

The Keneba study established procedures for the unique identification of each individual, enabling all the data linkages that are needed to exploit fully half a century of data on this Gambian population (Sear and others, 2000, and 2001; Rayco-Solon and others, 2004; Allal and others, 2004). In contrast with the Pholela study, the Keneba study is still ongoing and is the longest lasting DSS in Africa. It is however not as well known as more recently launched initiatives. The long delay in the dissemination of demographic findings from Keneba is probably due, in part, to the small size of the study population (about 1,000 people initially) and the consequent need for decades of accumulation of person years of observation before meaningful demographic analyses could be conducted. By 1975, a quarter of a century after the launch of the Keneba study, fewer than 2,000

person/years were accumulated on children under five years of age (estimates based on data produced by Rayco-Solon and others, 2004).

In 1962, Pierre Cantrelle, a French epidemiologist, launched a study to test the feasibility of reliably recording births, deaths, migration and marriages in the Niakhar and Paos-Koto districts of the Sine Saloum region in Senegal. This was the first African study where demographic objectives seemed clearly dominant. It was part of the newly born Senegalese State's effort to develop its vital registration system. During the first phase of the study (1962-1966), a total population of more than 50,000 was monitored in different areas. After a few years of trial, the decision was made to focus the research on the Niakhar district alone, hence the name of the study (Garenne and Cantrelle, 1991). According to Cantrelle, "only two published studies [of similar type, e.g. a DSS] were available when he started: the Yangtze River Valley study in China and the Guanabara study in Brazil. The Khanna study had just finished and had not yet been published while the Keneba study and the Pakistan Growth Experiment were underway" (Garenne and Cantrelle, 1991; p. 19).

These pioneers of the DSS approach in sub-Saharan Africa contributed to the science of demographic surveillance more generally. The success of the Niakhar study was translated into commitment to develop future initiatives. In the 24 years following the launch of the Niakhar study, five DSS studies were initiated, all of which were based in the Senegambia region of West Africa (Senegal, Gambia and Guinea-Bissau), except the short-lived Machakos study in Kenya (1974-1981). The launch of the Butajira study in Ethiopia in 1987 was the precursor to a rapid expansion of geographical coverage into Eastern and Southern Africa.

Table 1 provides a list of the ongoing programmes. It shows that about half the sites are concentrated in West Africa. The other half are distributed between Eastern and Southern Africa while Central Africa is virtually empty. Despite their high costs in human, material and financial resources, African DSS sites have been highly sustainable. Of the 31 sites that were launched since 1940, only 4 have ceased to exist. Currently, 27 sites are still functional.[2]

Table 1. Ongoing demographic surveillance sites in Africa

Site name	Country	Start date	Age in years	Popu-lation*	Lead institutions
West Africa				**907 800**	
Sapone	Burkina Faso	2005	2	100 000	MOH, CNRFP
Dodowa	Ghana	2005	2	98 200	MOH, Ghana Health Services
Kintampo	Ghana	2003	4	140 000	MOH, Kintampo Health Research Centre

.../

Table 1. *(Continued)*

Site name	Country	Start date	Age in years	Popu-lation*	Lead institutions
Ouagadougou	Burkina Faso	2001	6	5 500	University of Ouagadougou, UERD/ISSP
Navrongo	Ghana	1993	14	141 000	MOH, Navrongo Health Research Centre
Oubritenga	Burkina Faso	1993	14	175 000	MOH, CNRFP
Nouna	Burkina Faso	1992	15	55 000	Centre de Recherche en Sante de Nouna, U. of Heidelberg
Mlomp	Senegal	1985	22	7 600	Centre National de Recherche Scientifique, Musée de l'Homme, Ined
Farafenni	Gambia	1981	26	43 000	MRC-Gambia
Bandim	Guinea Bissau	1978	29	101 000	MOH, Danish Epidemiology Science Center
Bandafassi	Senegal	1970	37	10 500	Centre National de Recherche Scientifique, Musée de l'Homme, Ined
Niakhar	Senegal	1962	45	30 000	Institut pour la Recherche et le Developpemnt (IRD)
Keneba	Gambia	1950	57	1 000	MRC-Gambia
East Africa				**743 000**	
Iganga-Mayuge	Uganda	2005	2	62 000	Makerere University
Kilifi	Kenya	2003	4	225 000	KEMRI and Oxford University
Kisumu	Kenya	2001	6	135 000	CDC and KEMRI
Nairobi NUDSS	Kenya	2000	7	60 000	APHRC
Rufiji-Ifakara	Tanzania	1998	9	85 000	Ifakara Centre for Health Research and Development, former TEHIP
Ul/Kil-Ifakara	Tanzania	1996	11	65 000	Ifakara Centre for Health Research and Development
Magu	Tanzania	1994	13	28 000	TANESA program, Mwanza
Rakai	Uganda	1989	18	42 000	Makerere, Johns Hopkins and Columbia universities
Butajira	Ethiopia	1987	20	41 000	Unversity of Addis Ababa and University of Noumea
Southern Africa				**227 000**	
Karonga	Malawi	2002	5	30 000	London School of Hygiene & Tropical Medicine, Karonga Prevention Study
Hlabisa	South Africa	1999	8	85 000	ACDIS, University of Kwazulu Natal
Manhica	Mozambique	1996	11	37 000	MOH, Centre de Investigacao em Saude da Manhica
Dikgale	South Africa	1995	12	8 000	University of the North Digale Project
Agincourt	South Africa	1992	15	67 000	University of the Witswatersrand
Total Africa				**1 877 800**	

* Last population size known. Initial population size for Keneba is estimated at 1,000.

Table 2 shows a remarkable explosion of DSS sites in the early 1990s. The average number of new sites established every year increased consistently with time, from 0.1 in the 30 years following the launch of the Keneba study to 1.5 sites for the period starting in 2000. In addition, there was a marked increase in the size[3] of the populations being covered by these surveys. Table 2 shows a doubling of the average size of a DSS from 1990 onward. The size of the DSS populations has varied from 1,000 in Keneba (initial size in 1950) to 225,000 in Kilifi, Kenya.

Table 2. Evolution of demographic surveillance system site characteristics over time

Period	Number of new sites	New sites per year	Added population	Average population per site
2000-2006	9	1.5	855 700	95 078
1990-1999	10	1.0	746 000	74 600
1980-1989	4	0.4	133 600	33 400
1950-1979	4	0.1	142 500	35 625

This acceleration in DSS would not have happened without an increased demand for longitudinal community health research, but it was definitely rendered possible by technological breakthroughs that reduced the complexity of data management and the creation of the INDEPTH Network in the 1990s.

The complexity of data management and analysis has long represented a major challenge to the sustenance of DSS-based research. The initiation of field stations requires heavy investments in time, money and expertise for the development of software and a foolproof data management system. The development of the household registration system (HRS) and its subsequent dissemination across the world simplified the process dramatically. The relevant software was developed in collaboration with the Population Council with financial support from the Thrasher Research Fund and the Rockefeller Foundation. It was developed with the explicit intention of strengthening the capacity of developing countries to conduct demographic surveillance (Phillips and MacLeod, 1995).

HRS, which was inspired and shaped by the Matlab DSS, incorporates rigorous design principles, performs consistency checks and computes and reports key demographic indicators. It is easily adaptable to any longitudinal household study with minimal additional programming, and modifications do not require extensive knowledge of database management.

Generating such a complex system required a team of skilled computer scientists and demographers, in addition to field capacity (provided by the Navrongo Health Research Centre) in order to respond to queries and to advise on needed changes. The development of HRS constitutes a landmark in the history of longitudinal community studies, which broke the long-lasting problem of surveillance systems: the difficulties in data management and analysis. The mechanization of the complex logic of surveillance systems contributed significantly to reducing dependence on computer specialists, thereby reducing the cost of starting and sustaining DSS.

Since its adoption of HRS in 1993, the Navrongo Health Research Centre team played a key role in the dissemination and improvement of the software concerned. Many international workshops were organized to train interested parties in the use and adaptation of the software. Many visitors from other field sites spent some time in Navrongo to develop the hands-on expertise needed to implement an HRS-based solution for their sites. The software was freely downloadable from the Internet; technical advice was also readily available for its adaptation to local conditions.

This international role of the Navrongo Health Research Centre motivated the Rockefeller Foundation's decision to provide seed funding to the Centre to lead efforts to develop a network of longitudinal community studies. The birth of the INDEPTH Network in 1998 was a landmark event in the development of these studies in Africa. Not only did it provide a larger forum for the continued dissemination of HRS, but it also contributed to developing the software's capabilities further through confrontation with competing technology utilized at other sites and the creation of an INDEPTH reference data model[4] covering aspects that are common to all sites (Benzler, Herbst and MacLeod, 1998; Phillips and others, 2000; INDEPTH Network, 2002). The INDEPTH Network's technical working group took the technology of software design for DSS to new heights by broadening thinking beyond the dominating HRS (see for example Benzler and Clark, 2005; and Clark, 2006). Of the existing 27 sites, 9 of them were launched in the 8 years following the creation of the network!

A half century of contributions to science

Longitudinal community studies have made groundbreaking contributions to public health knowledge over the years. An ongoing effort to analyse these contribution has already identified more than 1,100 peer-reviewed publications resulting from these studies (Brosius and others, 2007). While awaiting an in-depth analysis of this mass of information in the next stages of the study, it is

worth pointing to a few contributions that would have been impossible without the rigour of the DSS approach.

One early contribution is the documentation of highly specific levels and patterns of mortality in childhood. The studies in the previously mentioned Senegambia region revealed extremely high levels of mortality, comparable to levels observed in seventeenth century Europe. A child mortality rate of about 500 per 1,000 live births (which meant that half of the newborn children did not reach their fifth birthday!) was observed in the early years of the Niakhar study. Furthermore, the age pattern of child mortality was very unusual. Mortality between exact ages 1 and 5 years was significantly higher than infant mortality, contrary to what is usually observed among populations outside of West Africa (Cantrelle, 1972). Exactly the same pattern was discovered by the Keneba study in neighbouring Gambia (McGregor, 1968, quoted by Cantrelle, 1972). Except in a few atypical years, the mortality rate between ages 1 and 5 years was consistently 50 per cent above the infant mortality rate in Niakhar from 1963 to 1990 (Delauney, 1999).

The DSS contribution to the characterization of mortality levels and patterns culminated with the development of authentic model life tables for sub-Saharan Africa (INDEPTH Network, 2004). These life tables, which condense the data collected in 19 DSS sites over 5 years, capture the diversity of the continent and reflect well the differential impact of the AIDS epidemic.

Studies of the relationships between the different components of population dynamics are impossible without detailed longitudinal data. The landmark article by Cantrelle and Leridon (1971), describing the relationships between fertility, mortality and breastfeeding, is one of the major contributions of DSS in its early phases of development in Africa.

Most of the initiatives launched since the 1990s were set up to study the impact of health interventions. This is an area where longitudinal community studies do not have much competition. The bibliometric analysis referred to previously (Brosius and others, 2007) found that 30 per cent of DSS-related publications are on health interventions. These studies provide insights into which population and health interventions work, why they work and how they work. Examples include the following: community health in Ghana (Phillips, Bawah and Biaka, 2006), integrated management of childhood illnesses in the United Republic of Tanzania (Schellenberg and others, 2004) and decentralized health financing in the United Republic of Tanzania (de Savigny and others, 2004).

Longitudinal studies also contributed to the scientific validation, through field trials of most health technologies in use today, including drugs, vaccines,

vitamins and insecticide-treated bednets (Aaby and others, 1986; Ross and others, 1995; Binka and others, 1996; Lengler and others, 1998).

Taken together, the ongoing longitudinal studies cover relatively well the diversity of situations in sub-Saharan Africa. A study of the environmental footprint of INDEPTH member sites in Africa concludes that these sites "provide a comprehensive coverage of much of the wide range of climates and environments" found across sub-Saharan Africa" and that "findings and relationships derived at DSS sites can be applied elsewhere within the same environmental class with some confidence." (Tatem and others, 2006). The initiation of additional sites in Nigeria and in central Africa would definitely provide better coverage.

Challenges and Africa's contribution to addressing them

Data management and processing

Without any doubt, the major challenge facing longitudinal community studies lies in the complexity of managing the large databases that such studies generate and stimulating the "production of science" by simplifying access to the data that are stored. Fortunately, as has been discussed in the preceding section, technological advances and their broad dissemination through cross-site collaboration are progressively lifting this major constraint.

The development of the INDEPTH reference data model is promoting common data management systems that will facilitate the construction of generic tools for data extraction and analysis, enabling scientists to work with complex databases just as they now work with commercial statistical packages.

Data analysis, utilization and sharing

African longitudinal community studies face one major and persistent criticism: the levels of data analysis and utilization to guide policy and practice are low. Data analysis is often limited to the minimum required to satisfy commitments to funding partners. Underanalysis is hampering the scientific community's ability to harness fully DSS data in order to provide the evidence base needed to improve public health. This unfortunate situation is due to many factors ranging from capacity shortages to ethical challenges arising from the need to protect the privacy of contributing populations.

The members of the INDEPTH Network understand technology's key role in addressing these challenges and is investing significant efforts in developing a platform that will further break down barriers between users of data and the mountain of DSS-generated information, stimulate cross-site comparative research

and facilitate access to the larger scientific community. These efforts on the technological front need to be backed by strategies to improve site leaders' willingness to share data.

Reducing cost without sacrificing quality

DSS will remain unattractive as long as data collection is seen as an expensive competitor to service delivery for scarce resources. Innovative ways of reducing the cost of data collection without jeopardizing data quality are urgently needed. Here, too, technology plays a key role. For example, handheld computers have become so powerful that they make it possible for the designers of computer systems to place some data management operations, such as automatic edits and consistency checks, in the hands of interviewers. By eliminating the cost and complexity of paper-based information transfer, DSS per capita operating costs can be reduced substantially.

While the use of handheld computers in DSS data collection is just beginning, one can foresee a future where very simple remote updating procedures would use mobile telephones as data entry devices or very wide area network cards as transmitters of information. These and other hardware innovations may transform DSS field operations by enabling key informants to signal DSS managers where interviewing should take place, thereby reducing significantly the number of unnecessary household visits. The experience of the Karonga DSS in Malawi suggests that innovative but less sophisticated use of the traditional key informant can go a long way in reducing costs without critically jeopardizing data quality (Jahn and others, 2007).

Research capacity

Another major challenge lies in the stringent human capacity requirement of DSS that is commensurate with the complex management and scientific tasks that are involved. Most African sites are unable to address effectively their capacity requirements. Because they are located typically in rural and remote areas, site leaders confront major difficulties in attracting and retaining local scientists. It is however clear that, because they entail long-term involvement with communities, longitudinal community studies have great potential for building both individual and institutional capacity for scientific inquiry and action. The experiences of Matlab in Bangladesh and Navrongo in Ghana provide vivid expressions of this potential and attest to the fact that, for greater effectiveness, capacity-building must be an explicit goal of the initiative.

Some principles of surveillance systems define contexts for capacity-building that other research strategies lack. The complexity of operations requires

career ladders, with entry-level interviewers, supervisors, managers and scientists at various levels of expertise. If the recruitment of talented young entry-level staff is combined with sustained career advancement strategies, longitudinal data systems can become engines for the development of scientific talent. This is particularly evident when sites are aligned with local and international university training programmes, and when degree-level training is augmented with site-based workshops and technical exchange. The fact that sites are engaged in critical policy and scientific issues can provide local scientists and support staff with opportunities for the dissemination of scientific information that other individuals lack. Exposure to international learning, combined with the challenges of operating sites, develops capacity, builds leadership and sustains African science in ways that other institutional arrangements cannot match.

Ethical challenges

Given the dire health situation in most of Africa, longitudinal community studies cannot afford the luxury of remaining purely scientific endeavours with no visible benefits for local communities, regional health administrations or national health programmes. In the absence of strategic planning about community engagement, research utilization and policy impact, there can be serious ethical tensions between science and service.

Furthermore, the need for unambiguous identification of each individual and household combined with the ubiquity of modern technology, makes it an obligation to develop ethical safeguards that ensure that the privacy and rights of respondents are fully protected.

Biomarkers may eventually replace identification numbers as the basis for data linkage, making DSS data models compatible with health information systems and providing direct links between service delivery information and DSS population registers. As these technologies are refined and applied, the potential scope for expanding the policy, administrative and research relevance of DSS data models will be enormous. Combining these benefits with ethical safeguards will represent a critical challenge for DSS planning in Africa.

Conclusions

It is fair to say that longitudinal community studies have improved our understanding of the ecology of health and disease, including the physical and social environment shaping human health in Africa. However, despite its palpable contributions to scientific knowledge, the DSS paradigm is often criticized for its complexity, cost and isolation from the policy processes. In Africa where these

studies are most needed, the shortage of technical human and financial resources has required sites to collaborate, share technology and pursue common strategies. As a consequence, African sites have been at the forefront of developing new technology for cross-site research, methods for reducing costs and general procedures for starting new DSS operations.

While this article is focused on health research, the usefulness of DSS as a research tool extends far beyond population and health. The DSS approach is valuable in all socio-economic studies requiring a measurement of change and its sources, including behavioural change, ageing and life cycle changes. Longitudinal community studies in Africa and elsewhere have the potential to contribute significantly to monitoring progress towards attainment of the Millennium Development Goals and particularly to measuring the impact of poverty reduction programmes, which are at the heart of development efforts. We hope that the INDEPTH Network will help to ripen this potential by sustaining and enhancing its efforts aimed at strengthening DSS sites throughout the world.

Endnotes

1. Three publications provide a comprehensive picture of this kind of study across the world: Das Gupta, 1997; INDEPTH Network, 2002; NRC, 2002.

2. The AMMP project in the United Republic of Tanzania was active in three areas but is counted here as one site. Its focus on adult mortality worked against the reliable recording of births.

3. Population size is important because the smaller is the size of a study population, the longer that study needs to continue in order to accumulate the number of person-years required for a reliable measurement of rare events.

4. "Data models" refer to predicted logical relationships that must be sustained despite any possible event that might be registered in a DSS database over time. For example, a pregnancy must have a woman associated with that event, she must be present as a member of the household where the event is registered and her age must fall within the reproductive years extending from menarche to menopause. Over time, women become ineligible for these events. Also, pregnancies must eventually end and other logic must prevail. All such logical specifications have a structure, order and predictability that permit software to identify errors and precisely define who is present, the days of risk associated with that presence and the events of interest that arise. Episodes of various kinds that members of a population experience are predictable with a data model, and changes in episode status are logically predictable as events unfold. The challenge in designing a DSS computer system is defining all such events, episodes and relationships. Tools that are now available to software developers permit the mechanization of computer logic once the underlying model is defined.

References

Aaby, P., J. Bukh, D. Kronborg, I.M. Lisel and M.C. De Silva (1986). "Vaccinated children get milder measles infection: A community study from Guinea-Bissau", *Journal of Infectious Diseases*; vol. 154, No. 5, pp. 858-863.

Allal, N., R. Sear, A.M. Prentice and R. Mace (2004). "An evolutionary model of stature, age at first birth, and reproductive success in Gambian women", *Proceedings of the Royal Society of London, Series B, Biological Sciences,* vol. 271, pp. 465–470.

Benzler, J. and S.J. Clark (2005). "Towards a unified timestamp with explicit precision", *Demographic Research*, vol. 12, No. 6, pp. 107-140.

Benzler, J., K. Herbst and B. MacLeod (1998). *A Data Model for Demographic Surveillance Systems,* <www.indepth-network.org/publications/indepth_ publications.htm>

Binka, F.N. and others (1996). "Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial", *Tropical Medicine and International Health*; vol. 1, No. 2, pp. 147-154.

Brosius, A. and others (2007). "The contribution of demographic surveillance systems: An indepth bibliometric analysis of publications from developing countries", paper presented to the PAA Annual Meeting, New York City, 2007.

Cantrelle, P. (1972). "Existe-t-il un type de mortalité tropicale?", *Afrique Médicale,* vol. 204, pp. 933-940.

Cantrelle P. and H. Leridon (1971). "Breast feeding, mortality in childhood and fertility in a rural zone of Senegal", *Population Studies-A Journal of Demography*; vol. 25, No. 3, pp. 505-533.

Clark, S. (2006). "A general temporal data model and the structured population event history register", *Demographic Research*, vol. 15, No. 7, pp.181-252.

Das Gupta, M., P. Aaby, M. Garenne and G. Pison (1997). *Prospective Community Studies in Developing Countries*, Oxford, Initrd Kingdom, Clarendon Press.

Delauney, V. (1999). *La baisse de la mortalité des enfants en milieu rural Sénégalais: l'exemple de la zone de Niakhar,* in Adiamagbo, A., G. Agnes and K. Nguessan, eds., *Santé de la Mère et de l'Enfant: Exemples Africains*, Paris, Institut de recherche pour le developpment, pp. 131-156.

de Savigny, D., H. Kasale, C. Mbuya and G. Reid (2004). *Fixing Health Systems*, Ottawa, International Development Research Centre.

Garenne, M. and Pierre Cantrelle (1991). "Three decades of research on population and health: The ORSTOM experience in rural Senegal, 1962-1991", paper presented at the IUSSP Seminar on Sociocultural Determinants of Morbidity and Mortality in Developing Countries: The Role of Longitudinal Studies, Saly Portudal, Senegal, 7-11 October.

INDEPTH Network (2002). *Population and Health in Developing Countries, Volume 1: Population, Health and Survival at INDEPTH sites,* Ottawa, International Development Research Centre.

_____ (2004). *INDEPTH Model Life Tables for sub-Saharan Africa*, Alderskot, United Kingdom, Ashgate Publishing Company.

Jamison, D.T. and others (2006). *Priorities in Health*, Washington, D.C., World Bank.

Jahn, A. and others (2007). "Evaluation of a village-informant driven demographic surveillance system in Karonga, northern Malawi", *Demographic Research,* vol. 16, No. 8, pp. 219-248.

Lengeler, C. and others (1998). "Relative versus absolute risk of dying reduction after using insecticide-treated nets for malaria control in Africa", *Tropical Medicine and International Health,* vol. 3, No. 4, pp. 286-290.

McGregor, I.A. and D.A. Smith (1952). "A health, nutrition, and parasitological survey in a rural village (Keneba) in West Kiang, Gambia", *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 46, No. 4, pp. 403-427.

National Research Council (2002). *Leveraging Longitudinal Data in Developing Countries: Report of a Workshop,* Valerie L. Durrant and Jane Menken, eds., Washington, D.C., National Academy Press.

Ngom, P. and others (2001). Demographic surveillance and health equity in sub-Saharan Africa, *Health Policy and Planning*, vol. 16, No. 7, pp. 337-344.

Omran, A.R. (1974). "Population epidemiology: an emerging field of inquiry for population and health students", *American Journal of Public Health*, vol. 64, No. 7, pp. 674-679.

Phillips, J.F. (2003). "Demographic surveillance systems", in Paul Demeny and Geoffrey McNicoll, eds., *The Encyclopedia of Population,* London, MacMillan, pp. 204-206.

_____ and B. MacLeod (1995). *The Rapid Transfer of Demographic Surveillance Systems with Automated Software Generation Technology: A Final Report to the Thrasher Fund.*

_____, A. Bawah and F.N. Binka (2006). "Accelerating reproductive and child health program impact with community-based services: The Navrongo experiment in Ghana", *Bulletin of the World Health Organization,* vol. 84, pp. 949-955.

Rayco-Solon, P., S.E. Moore, A.J. Fulford and A.M. Prentice (2004). "Fifty year mortality trends in three rural African villages", *Tropical Medicine and International Health*, vol. 9, No. 11, pp. 1151-1160.

Ross, D.A., B. Kirkwood, F.N. Binka, P. Arthur, N. Dollimore, S.S. Morris, R.P. Shias, J.O. Gyapony and P.G. Smith (1995). "Child morbidity and mortality following vitamin A supplementation in Ghana: Time since dosing, number of doses, and time of year", *American Journal of Public Health* vol. 85, No. 9, pp. 1246-1251.

Schellenberg, J.A. and others (2004). "Effectiveness and cost of facility-based integrated management of childhood illness (IMCI) in Tanzania", *The Lancet*, vol. 364, pp. 1583-1594.

Sear, R., R. Mace and I.A. McGregor (2000). "Maternal grandmothers improve the nutritional status and survival of children in rural Gambia", *Proceedings of the Royal Society, Series B, Biological Sciences,* vol. 267, pp. 1641-1647.

_____, D. Shanley, I.A. McGregor and R. Mace (2001). "The fitness of twin mothers: Evidence from rural Gambia", *Journal of Evolutionary Biology*, vol. 14, pp. 433-443.

Taten, A.J., R.W. Snow and S.I. Hay (2006). "Mapping the environmental coverage of the INDEPTH demographic surveillance system network in rural Africa", *Tropical Medicine and International Health*, vol. 11, No. 8, pp. 1318-1326.

The Cebu Longitudinal Health and Nutrition Survey: Two Decades Later

Many valuable lessons can be drawn from the 23-year history of the Cebu Longitudinal Health and Nutrition Survey, the longest health and nutrition panel study in the Philippines and among the longest such studies in Asia. The survey offers an example for other developing country settings.

By A.B. Feranil, S.A. Gultiano and L.S. Adair*

This article discusses the 23-year history of the Cebu Longitudinal Health and Nutrition Survey (CLHNS), which was started in 1983 in collaboration with the Carolina Population Center of the University of North Carolina at Chapel Hill (CPC-UNC), the Office of Population Studies of the University of San Carlos (OPS-USC) and the Nutrition Center of the Philippines (NCP). After briefly presenting information on its origins and development in this introductory section,

* Alan B. Feranil and Socorro A. Gultiano are Senior Research Associates of the University of San Carlos, Office of Population Studies Foundation, Philippines, and Linda S. Adair is a Fellow of the Carolina Population Center of the University of North Carolina at Chapel Hill in the United States of America. E-mail: alanferanil_2000@yahoo.com.

the methodologies and strategies used in the research and some of its major findings and contributions are discussed, before analysing the challenges and lessons learned from the study. The article ends by identifying some best practices and offering some recommendations for other longitudinal studies.

The origin of CLHNS can be traced to the controversy surrounding the determinants and consequences of infant feeding in the early 1980s, which prompted several international companies producing infant formula to approach Barry Popkin, a noted expert on nutrition, asking him to develop a long-term research project to explore some of the unresolved issues concerning this matter (Adair and Popkin, 2001). Popkin formed a research group which included economists John Akin and David Guilkey from CPC-UNC, demographer Wilhelm Flieger from OPS-USC and Florentino Solon, Director of the Nutrition Center of the Philippines. Metro Cebu was chosen as the research site because of its environmental and socio-economic diversity [1](Adair and Popkin, 2001).

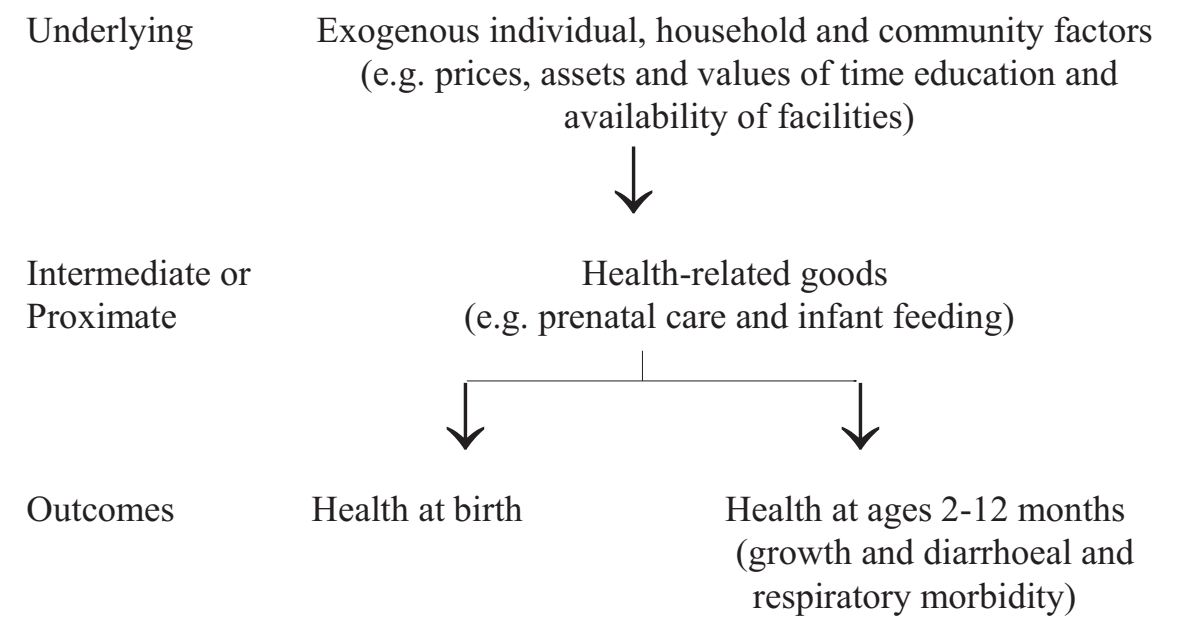
The study started with the objective of understanding infant feeding practices, the determinants and consequences of infant feeding and maternal and child health practices, and the effects of ecological factors on birth and health outcomes. The initial years of the project also investigated how infant feeding decisions were made, the frequency of feeding events using milk and non-milk feeding and the effects of feeding patterns on mothers and households (OPS-USC, CPC-UNC and NCP, 1989).

CLHNS was one of the first large-scale population-based surveys designed with a conceptual framework that was similar to what would eventually become known as Mosley and Chen's (1984) framework for child survival. The framework argues that individual household and community variables affect a set of proximate health behaviours (infant feeding and use of preventive health care), which in turn influence health outcomes including morbidity and growth (Adair and Popkin, 2001). As shown in figure 1, the framework used by the study examines the underlying individual and community factors (prices, assets, values of time, education and availability of facilities), intermediate or proximate factors (health-related "goods" such as prenatal care and infant feeding) affecting health outcomes at birth (health at birth, birth weight and gestational age) and health at ages 2-12 months (growth and diarrhoeal and respiratory morbidity).

Although the CLHNS was originally intended for a two-year follow-up of infants, subsequent surveys were conducted in 1991-1992, 1994-1995, 1998-1999, 2002 and 2005 on these cohorts of mothers and children to pursue a variety of enquiries on different aspects of the mother and child's development to be described in greater detail in a later section of this article. For each new research

objective, specialized modules were developed and added to the core/basic modules retained from earlier rounds, building on the large information database of CLHNS. The evolutionary nature of CLHNS is what makes the study unique.

Figure 1. Initial research framework relating underlying intermediate and outcome variables



Source: OPS-USC, CPC-UNC, NCP (1989), Cebu Study Team (1991).

Methodology, data and attrition

CLHNS used a stratified single-stage cluster-sampling procedure to select 33 *barangays*, (the Filipino term for a village or district), indicating the smaller administrative units in the Philippines, of which 17 were urban[2] and 16 rural in Metro Cebu; a cohort of pregnant women from those locations were to be included in the study. Stratification was by urban or rural residence (defined by the census) and the *barangay* served as the cluster. Infants born between 1 May 1983 and 30 April 1984 and their mothers served as the baseline sample of 3,327 mother-and-child pairs. A 12-month birth cohort was recruited to account for seasonality as a possible determinant of child growth, morbidity, mortality, maternal fertility and nutritional patterns (Adair and Popkin, 2001). Data were collected on all the infants born in the sample *barangays* and on their mothers. Interviews and measurements were undertaken immediately after each birth, then at two-month intervals for 24 months. Data on the mother and child within a day or two of the bimonthly anniversary of each child's birth over a two-year period were

collected to capture the sequence of feeding events, the dynamics of breastfeeding, feeding behaviour and the factors affecting feeding decisions at each point in time.

Several strategies were used to collect information and ensure data quality. These covered different stages of the project from sampling and training to data collection. To start, a household census was undertaken to identify pregnant women to be included in the study. A house-to-house survey was conducted in the 33 sample *barangays* in 1982 and 1983. All women who were in their third trimester of pregnancy at that time and who subsequently gave birth between 1 May 1983 and 30 April 1984 were recruited for the study (Gultiano, 1999a).

Community preparation was made with visits to provincial, municipal and *barangay* officials prior to the survey to inform officials of the planned project and obtain their permission to conduct the study in their jurisdiction, and to get their support and acceptance for the survey to be conducted in their community. Meetings informed the populace or community concerned of the study to be undertaken. Official permits were also obtained from administrative officials in all cities and municipalities and from health officials and military authorities (OPS-USC, CPC-UNC and NCP, 1989).

A *barangay* reporting system was established to obtain reliable and speedy information about infant births and infant deaths that occurred within the sample *barangays*. One or two residents in the *barangay* were hired as regular reporters to provide the needed information and updates on the sample children directly to the survey teams or the office.

Traditional health workers and midwives were enlisted to report all the births they attended in the sample *barangays*. Incentives were given for incidents reported within three days, considered the ideal time to gather information since data on the infant food intake could also be obtained, and even more incentives were given if births were reported immediately and infants weighed using the scale provided by the project. Birth interviews taken more than two weeks after the birth were ruled as unreliable because of recall problems and measurement bias. These midwives and traditional healers were also asked to report women who became pregnant after the household listing in March 1983 and those who were missed during the operations. These reports were verified by field personnel; this strategy resulted in 430 women who were either newly pregnant or had been overlooked (OPS-USC, CPC-UNC and NCP, 1989).

Intensive training of field personnel ensured the quality of data collection. Seasoned OPS staff served as field supervisors. Nurses were hired for the administration of gestational age tests. Field supervisors and staff were also

trained to obtain reliable information on infant morbidity, especially diarrhoea, measles and respiratory infections, anthropometric measurements, gestational age and dietary intake recall under the supervision of health, child and nutrition experts from the Cebu Institute of Medicine and University of the Philippines at Los Baños (OPS-USC, CPC-UNC and NCP, 1989).

A permanent team was assigned in designated sample *barangays* for the first two years of the study. Data collection was done on a bimonthly basis. This strategy distributed the work evenly across rural, urban and island sample *barangays*, and facilitated timely data collection by the survey team members familiar with the survey site whose presence in the community was already accepted. To avoid bias in reporting or documentation that may have arisen owing to the familiarity of the interviewer with the respondent's history, assignments among field interviewers were made at random. This strategy also proved useful in validating previously collected data and identifying interviewers who were not doing their job efficiently.

Regular checks were in place to ensure good quality data. Apart from the intensive training of interviewers, interobserver reliability was periodically assessed. All completed questionnaires were edited for consistency in the field and in the office. Data entry included range checks; questionable responses were sent back for a reinterview.

The construction of a data entry software program suitable for longitudinal data and creation of unique identification cards was done to suit the longitudinal, hierarchical data structure and complexity of data, which included special modules in different survey rounds. Unique identification numbers were generated for each mother and child in order to facilitate the merging of data files from different survey rounds.

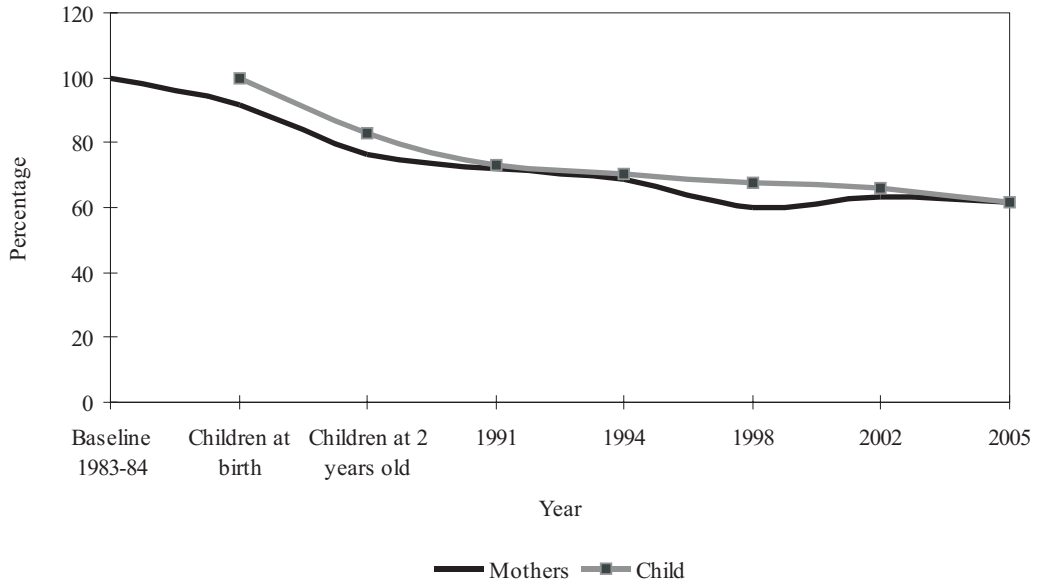
The study ensured that the conduct of the study was ethically sound. Apart from seeking clearance from the institutional review board, the research team underwent training in the ethical conduct of research. Verbal or written informed consent was obtained before proceeding with the interviews or data collection. Written informed consent was obtained when biological specimens (blood and saliva) were collected in later surveys. For sensitive topics, a sealed envelope was used to respect the privacy of the respondents and enhance the response rates, such as on the study of pubertal development when adolescents were asked to self-evaluate their status relative to a set of line drawings of the different stages of puberty.

A persistent follow-up of sample mothers and children was done using all possible means. Visits were scheduled to accommodate the work and school

schedules of the respondents, including, if needed, conducting interviews at their schools or places of work. Tokens and "bother fees" were given to the participating mothers to compensate for the opportunity costs or the income they would have otherwise earned if they had been working.

This careful planning helped to address the issue of attrition, a common problem in longitudinal studies. Although not all respondents could be retained (see figure 2), after more than 20 years, the retention rate is still comparable to, if not better than, other cohort studies worldwide. In 2005, of the original 3,327 mothers 2,080 were still participating in the survey. Attrition among children was similar to that of the mothers. From an original cohort of 3,080 single live-birth children, only 1,903 remained in 2005, representing the same percentage as that of the mothers (Gultiano, 2006).

Figure 2. Percentage of mothers and children included in the rounds of the Cebu Longitudinal Health and Nutrition Survey



The leading cause of attrition has been migration outside of Metro Cebu, accounting for a loss of about 3 in 10 mothers and a fourth of the children in 2002 (Perez, 2003). Deaths accounted for a very small percentage of the loss of sample mothers and children. As expected, the death of children was higher in the earlier ages up to the first year of life (Perez, 2003). Refusals for interview were rather low across the years, ranging from less than 1 per cent to about 2 per cent for both mothers and children (Perez, 2003). Because of inevitable attrition, the latest sample of children in 2005 is considered to be biased towards rural and poor households with fewer modern amenities or assets and less educated parents (Perez and Cinco, 2007; Bas, 2007, Carba and Tan, 2007). Another noted bias is towards

older mothers (Adair and others, 1997). It is therefore important to test for selection biases in the analysis of more recent data.

Wealth of information

CLHNS data are rich and varied, covering three generations in more than two decades. The data and measurements included in the successive surveys were tailored to serve the objectives of each survey. The baseline, birth interview survey and the periodic surveys until each child was two years old were meant to measure the determinants of infant feeding and feeding practices and determine the influence of ecological factors on health outcomes (birth weight, mortality and infections during the first two years of life). The baseline survey (in the sixth or seventh month of pregnancy) focused on maternal health behaviour, diet (assessed using a 24-hour recall) and health status, and the household's socio-economic environment. Anthropometric measurements, particularly height, weight and arm circumference, were taken by trained field personnel. Key informants have provided information on *barangay* characteristics, educational facilities, social services, public utilities, *barangay* organizations, commercial establishments and the price of milk, infant foods and other foods and commodities in each community.

The birth interviews have included questions on delivery practices, and health and infant feeding during the first days of the infant's life. Infant weights were recorded by health personnel or traditional health providers, and then weight and length were measured by project interviewers. Gestational age was assessed using the Ballard assessment method,[3] for all infants born weighing less than 2.5 kg, or whose mothers had complications during pregnancy or whose gestational age could not be obtained from indicators normally supplied by the mothers (OPS-USC, CPC-UNC and NCP, 1989). Bimonthly surveys provided detailed information on infant and maternal diet, mother and infant anthropometry and morbidity, lactational patterns, physical activity and nutritional status, socio-economic household changes, health practices and behaviour, family planning and other child-spacing factors. Special modules were also administered during these rounds to assess water usage and quality, the incidence of diarrhoea and the influence of media advertising on breastfeeding and the feeding of infants. Community-level information included a survey of health professionals' knowledge, attitudes and practices, and community characteristics (taken biannually), food markets (bimonthly), water quality (quarterly) and family planning facilities and personnel (twice) (OPS-USC, CPC-UNC, NCP, 1989).

The specialized follow-ups produced a wealth of information on related topics. The first follow-up of the mother-and-child pairs was undertaken in the

period 1991-1992, when the children were eight years old, to follow the growth, intellectual development and early schooling status of the offspring. The 1991-1992 survey included the child's complete schooling history and administration of non-verbal intelligence tests and the measurement of the next youngest sibling's anthropometrics. Subsequent surveys were conducted in 1994-1995, 1998-1999, 2002 and 2005, with each survey looking at a different perspective of the mother and child's development. The 1994-1995 survey focused on women's status, family planning use and labour force participation of the mothers and the relationship between school achievement and nutritional status of the children (Gultiano, 2006). A module on household decision-making included questions designed to measure autonomy in several realms. Questions concerning the mothers' aspirations for their children were also asked during this round. Detailed in-depth interviews on a subset of 60 mothers were undertaken to explore how family planning decisions were made, the consequences of family planning use, the various aspects of their lives involving work, autonomy, status and aspirations (Gultiano, 1999b; Avila and Wong, 2001). Achievement tests in mathematics and in the English and Cebuano languages were administered to children who were then 10-11 years old.

The 1998-1999 survey round focused on adolescent reproductive health and sexual behaviour. During this round, the sample children were interviewed for the first time. Apart from the data collected in earlier rounds, adolescents were asked additional information about their physical development, using drawings for sensitive questions; work history; pregnancy history; breastfeeding practices; marital status; marriage and cohabitation; sexual relations; reproductive health; family, peer and media influence; risky behaviour; communication with parents; and daily activities. Blood pressure readings were also taken on all respondents Blood and saliva samples were taken from a selected subset of 600 adolescents, including all who were small at birth (weighing less than 2.6 kg), and a random sample of those who weighed more than 2.6 kg. Samples were analysed for hemoglobin, lipids, glucose, insulin, immune and thymic function and blood type. Saliva samples were analysed for cortisol levels.

For the mothers, complete reproductive histories added information on menopause; information was also collected to provide baseline measures for future studies of ageing, including a cognitive function memory test and questions about activities of daily living. Mothers were also asked questions reflecting their psychological health and communication with their daughters or sons. During this round, in-depth interviews with a subsample of fathers were undertaken to determine their views on family planning and women's work.

The more recent 2002 and 2005 survey rounds examined how childhood health and nutritional status affected educational attainment work patterns and the wages of young adults. The 2002 survey included a separate module on intimate partner violence/aggression, decision-making, financial management and parental status and body image and the start of the life-history matrix.[4] Additional anthropometrics included waist circumference and subscapular skinfold thickness. During this round, both the adult children and their spouses were included in the survey because of the interest in the domestic behaviour of wives and their husbands. The spouses were asked the same questions as the sample adult children. Modules on entertainment facilities, major business establishments, job availability and a special module on government financing and resources were also included in this round.

The most recent survey (2005) added interviews with husbands of the mothers, and information about the children born to index children, thus adding a third generation to the study. Questions about perceived stress were asked of mothers and their husbands during this round. Fasting blood samples were taken from the mothers and their children, now young adults. Husbands were interviewed using the same survey forms, but they provided no blood samples. In addition to the storage of plasma for assessment of a wide range of chronic disease biomarkers, DNA was extracted for future genetic analysis.

Major findings and contributions

The CLHNS data, which are publicly available, have enabled many scientists to conduct demographic, health, social and psychological research. Although the results relate to the Philippine context, some of the findings also have relevance for other developing countries. An exhaustive description is beyond the scope of this article, but a few of the main results are briefly mentioned to provide an impression of the relevance of CLHNS for knowledge, policy and action.

CLHNS has provided evidence about the importance of intrauterine development and early child development, especially in the first two years of life, and the implications of these aspects. Studies on prenatal care have shown the positive impacts on the health of infants and mothers.[5] These were complemented by studies on low birth weight and stunting and their respective impacts in childhood. Illness during early life was shown to affect growth as well as academic performance in late childhood. The link between intrauterine growth and consequences later in life has also been explored using the CLHNS dataset supporting the foetal programming or the development origins of a health and development hypothesis.

CLHNS has also yielded several studies on the positive effects of breastfeeding; the negative effects of breastfeeding, particularly on the mothers; the factors influencing breastfeeding; and the effects of breastfeeding on amenorrhoea. Studies on weaning focused on the effects of early weaning, infant feeding practices, use of breastmillk substitutes and supplementation.

The wide range of data collected in CLHNS from the 1980s to 2005 on the mothers, their children and more recently on their grandchildren provide vast opportunities for intergenerational studies that explore the influence of nutrition on the growth of children and the subsequent generation and parent-child studies from biological and social perspectives. The data also make possible the investigation of epidemiological and nutrition transitions in a developing country. Metro Cebu provides an excellent backdrop for these analyses, considering that it has experienced rapid modernization and development in the past several decades. Cebu Province has been experiencing the double burden of both infectious diseases and degenerative disorders (Quiza, 2006). Similarly, the population offers examples of both undernutrition and overnutrition (obesity and overweight) as observed in the recent 2003 National Nutrition Survey (Food and Nutrition Research Institute, 2006).

CLHNS data have been used to examine nutritional transition, maternal education and employment, human capital formation and childbearing, and to explore issues related to gender, adolescent and reproductive health, domestic violence and family planning.

Besides these content topics, a number of methodological issues could also be explored, including the effect of endogeneity and the use of structured questionnaires versus in-depth interviews.

These and many other outputs have brought recognition to CLHNS as one of the prime birth cohorts in a developing country, as reflected also in the inclusion of CLHNS in the COHORTS Study Group, a collaboration of research efforts from five sites, namely Brazil, Guatemala, India, the Philippines (Cebu) and South Africa, funded by the Wellcome Trust. Among the many contributions of CLHNS in building knowledge and capacity is the use of its findings in continuing education programmes of the Philippine Pediatrics Society, the Philippine Obstetrics and Gynecology Society and the Philippine Society for Nutrition and Dietetics and in academic publications. Its data have been elaborated in more than 100 publications and in numerous dissertations and conference papers. Most importantly, the results of CLHNS have had impacts on policies and interventions. According to Adair and Popkin (2001), the first World Bank strategy on health financing and environmental health policy used findings drawn

from the Cebu study. Likewise, the policy of UNICEF on the ingestion of breastmilk substitutes during infancy and the Asian Development Bank's programmes on the early health development of children have used findings from CLHNS.

Challenges and lessons

Several challenges were encountered in the conduct of CLHNS. These included ensuring the follow-up of the sample mothers and children, having a dedicated team of researchers, ensuring good data quality and database management, and securing funds for subsequent activities.

Following the sample mothers and children was done using all available sources of information (past and current, including that provided by their neighbours); however, resources were inadequate to follow respondents moving outside of Metro Cebu. Ethical questions also had to be taken into consideration, especially in studying populations in which severe health outcomes were being measured. Interviewers were trained to refer individuals in need of mental or physical health care, even if doing so implied losing a case. As mentioned previously, the survey staff had to go to great lengths to track those remaining within the boundaries of the survey in order to reduce attrition to a minimum.

In addition to retaining the respondents, it was a challenging task to retain the staff and maintain a dedicated team of researchers to follow up the cohorts owing to competitive job offers. In trying not to lose trained survey personnel who were familiar with CLHNS, benefits and good compensation were provided, although job security could not be guaranteed.

In the face of staff turn-over, high-quality training, technological development and maintenance of survey protocols were employed to ensure high-quality comparable data over the years. This was not an easy task, as consistency checks for each survey round and across survey rounds are complex. Managing data files for the longitudinal data was a major concern in view of the advancements made in computer technology over time and the volume of information to be processed in addition to constructing longitudinal data files that would be comparable across many survey rounds. Although data quality can be safeguarded with proper checks (editing in the field and office, and by machine), the employment of competent human resources to handle quality control proved expensive. These costs warranted the high quality of CLHNS data when compared with what is available from many developing country surveys, but they put pressure on fund-raising efforts.

Securing funds to continue with CLHNS was time-consuming, particularly in a climate of increased competition for fewer resources. By diversifying the study and defining a wide range of important health and social issues that could be uniquely addressed by CLHNS data, the authors have been quite successful in obtaining funding from a wide range of sources, including government, international organizations, research foundations, universities, industry and the United States National Institutes of Health, among other sources.

In addressing these and many other challenges for more than two decades, valuable lessons have been learned in the conduct of a longitudinal study, from which the following best practices could be identified:

- Documenting the procedures from each survey round is essential for creating user-friendly data sets. With proper and careful documentation of the developments and uniqueness included in each survey round, users can become aware of the history to guide them in understanding the data set.
- Maintaining a database to store the information from the different survey rounds is an important function as such a database is a major asset that needs to be supported by the development of computer programs which enable efficient data encoding, editing, archiving and use of the longitudinal data across surveys.
- Tracking or surveillance of the cohort of mothers and children ought to be carefully planned in order to know their whereabouts, the major events in their lives (e.g. graduation, employment, marriage, new births) and to update some measures. The master list of respondents should be frequently updated to facilitate the follow-up of respondents for the subsequent survey rounds. Tracking is very important in order to document the many events happening in the transition from adolescence to adulthood; it also enhances the rates of follow-up.
- Hiring a team of competent and dedicated researchers is critical to the success of the longitudinal study. Having them well trained and equipped with the skills to measure the information is a must for ensuring good data quality. Conducting standardized training programmes across surveys minimizes intra- and inter- interviewer bias. Although CLHNS has been fortunate to have some interviewers since the start, new interviewers had to be trained and undergo the same training process as the others.
- Enlisting of qualified local residents to serve as part of the survey team is an efficient way of getting timely information. Having local reporters to

report the birth of a child facilitated the entry of the survey team and the gathering of information within the first few days after the birth, thus minimizing recall problems. Enlisting traditional health providers as part of the research team also proved to be beneficial since this step assured community acceptance of the survey team.

- Apart from field, office and machine editing to ensure the quality of the data, monitoring the interviewers ought to be carried out to prevent erroneous data reporting. Although problems of faulty reporting of data may have been avoided with the imposition of allowance deductions for erroneous data, having the same interviewers returning to the participant may lead to some bias in reporting. To avoid behavioural contamination resulting from frequent follow-ups by the same survey team in the same sample *barangay*, starting in 1991 the interviewers were randomly assigned to the mothers, thus ensuring that the interviewers would not interview the same mothers.

- Establishing good rapport with the community, particularly the elders and natural leaders who influence the other community members, facilitates cooperation, which is key to the success of the study. On the down side, close interaction can be a taxing experience for some interviewers. In some cases, interviewers were asked for financial assistance to support the hospitalization or burial of a family member; in most of these cases, interviewers made contributions in their personal capacity.

- In ensuring long-lasting participation, it is important to inform respondents of the study and to be sensitive to their beliefs, customs and perceptions so that intrusion into the privacy or personal lives of the respondents would be acceptable, and survey operations would not be stifled. For example, during field tests of the anthropometric equipment, the survey team found that some participants perceived the measuring activity to be similar to measuring someone for a coffin; therefore, they adjusted the infantometer by removing its sides and painting it with lively colours to dispel such notions.

- Explaining the purpose of the study is crucial in getting the study participants' cooperation. Although some mothers had misperceptions and expressed anxiety when a test (Ballard assessment method) for determining the gestational maturity of newborn infants was administered, an explanation of the reasons for the measurements and the purpose of the study appeased them, after which they cooperated. Refusal of respondents to join the study can be best handled by providing them

with information on the benefits of the study. Providing a "bother fee" or token payment may also be useful in maintaining their cooperation.

- Engaging the community in the study can be fostered by devising ways for "giving back to the community". By informing the respondents of the study results through informal conversations or tangible forms (such as information brochures in vernacular languages, growth charts and blood readings) study participants were made aware of the relevance of the study to their lives and those of other people.

- Preserving continuity by maintaining questions in different survey rounds. In cases where questions may be flawed or inadequate, better questions can be incorporated in subsequent rounds. The analysis of CLHNS also paved the way for learning what to include and improve in subsequent rounds of the survey.

- Eliciting institutional collaboration among institutions plays a major role in helping to continue the cohort study. With support, training and collaboration among institutions, studies on CLHNS still continue.

- Observing proper research ethics is necessary; this involves maintaining transparency about the study. Informed consent is to be sought and the decision of subjects to participate or not participate in the study or submit specimens has to be respected. In sharing information, the anonymity and confidentiality of personal information should be ensured. With these practices and the constant assurance of confidentiality, participants appreciate the value of the research and their participation.

- Disseminating results to major stakeholders (such as funding organizations, academia, and governmental and non-governmental organizations) and the making of survey data available are vital for building knowledge and allowing for the provision of inputs to policy measures and interventions.

In conclusion, based on the CLHNS experience and summing up the lessons learned and the good practices identified, the authors would like to make the following recommendations. First, future endeavours on longitudinal studies should be culture- and people-sensitive so that the survey is understood by, and acceptable to, the participants, and beliefs and misconceptions of the sample population can be addressed early during the survey. Second, there is a need for transparency about the study, ensuring informed consent, respecting the rights and decisions of people to participate or not participate in the study or to submit

specimens. Third, there is a need to return to the people and obtain feedback on the results of the study either formally or informally so that they will appreciate the value of the research and their participation in it. Fourth, the different processes and experiences of different survey rounds should be documented, archived and shared with other users so that these can serve as learning experiences for other researchers. Fifth and lastly, it is necessary to disseminate and popularize the technical findings of the studies among key stakeholders so that the findings can serve their purpose and be utilized as inputs to policy measures and interventions, future research efforts and advancing scientific pursuits.

Acknowledgements

The authors would like to acknowledge with gratitude the contributions of Tita Lorna Perez, Isabelita Bas, Marilyn Cinco, Delia Carba, Nenette Lim and Josephine Avila of the University of San Carlos-Office of Population Studies Foundation in providing information on the history of CLHNS and insights into the conduct of the surveys. They would also like to thank their sponsors and donors who have helped them through the years, who believed in the value of this longitudinal study and ensured its continuation. Most of all, they are grateful to their collaborators who have contributed to CLHNS.

Endnotes

1. Metro Cebu, the second major metropolitan area in the Philippines, has a population of 3.3 million as of the last census (2000). Cebu Province provides a good backdrop for the conduct of socio-economic, demographic and health surveys since it is an island province and one of the fastest growing metropolitan areas in the country. The social, economic and health changes that the country has been experiencing are reflected in Cebu; the rapid epidemiological and nutritional transitions happening in Cebu typify those of any developing country in Asia (Adair and Popkin, 2001).

2. At the start of CLHNS there were 243 *barangays* in Metro Cebu, of which 155 were urban and 88 were classified as rural by the National Statistics Office. With the revised classification of the Office of Population, according to more stringent socio-economic criteria for the purposes of sample stratification, only 95 *barangays* were considered urban while 148 remained rural. Drawing separately from the two clusters, 17 urban *barangays* and 16 *barangays* were randomly drawn. There was an oversampling of the urban *barangay,* with the assumption that health care and infant feeding patterns are more diverse in the urban setting than in rural ones (Gultiano, 1999a).

3. Trained nurses were employed to administer the Dubowitz-Ballard test of gestational maturity (OPS-USC, CPC-UNC and NCP, 1989).

4. The life-history matrix is an instrument for recording and sequencing various events in a person's life. The matrix in the study included nine major aspects of the child's life, namely education, occupation, residence, romantic relationships, sexual experience, marriage, pregnancy, family planning and major illnesses. The matrix shows how changes in one event or milestone relates to changes in the other events or milestones.

References

Adair, L.S. and others (1997). "Cebu Longitudinal Health and Nutrition Study: Follow up study final report", Family Health International, North Carolina (unpublished).

_____ and B.M. Popkin (2001). "The Cebu Longitudinal Health and Nutrition Survey: History and major contributions of the project", *Philippine Quarterly of Culture and Society*, vol. 29, pp. 5-37.

Avila J.L. and E. Wong (2001). "When fate and husbands prevail: Dynamics of decisions and behavior regarding family planning in Cebu, Philippines", *Philippine Quarterly of Culture and Society,* vol. 29, Nos. 1-2, pp. 79-104.

Bas, I. (2007). "Childhood stunting and schooling attainment of Filipino young adults", poster presented during the 2007 Annual Meeting of the Population Association of America, New York, 29 March.

Carba, D. and V. Tan (2007). "Is early childhood height for age associated with the work status of young adults in Cebu Philippines?" poster presented during the 2007 Annual Meeting of the Population Association of America, New York, 29 March.

Cebu Study Team (1991). "Underlying and proximate determinants of child health: The Cebu Longitudinal Health and Nutrition Study", *American Journal of Epidemiology,* vol. 133, No. 2, pp. 185-201.

Food and Nutrition Research Institute (2006). *Philippine Nutrition Facts and Figures 2003,* Bicutan, Tagig, Food and Nutrition Research Institute.

Gultiano, S. (1999a). "A longitudinal perspective of mothers' employment in Metro Cebu the Philippines", doctoral thesis (demography) submitted to Australian National University (unpublished).

_____ (1999b). "Husbands' views on family planning and labor force participation of wives in Metro Cebu, the Philippines, *Philippine Quarterly of Culture and Society,* vol. 27, Nos. 3-4, pp. 133-160.

_____ (2006). The Cebu Longitudinal Health and Nutrition Survey (CLHNS): A multipurpose study of mother-child pairs in the Philippines (a cohort or panel study), Powerpoint presentation.

Mosley, J. and L. Chen (1984). "Analytical framework for the study of child survival in developing countries", *Population and Development Review*, vol. 10 (Supplement), pp. 25-45.

Office of Population Studies, University of San Carlos (OPS-USC), Carolina Population Center University of North Carolina (CPC-UNC) and Nutrition Center of the Philippines (NCP) (1989). "The Cebu Longitudinal Health and Nutrition Survey: Survey procedures and survey instruments" (unpublished).

Perez, L. (2003). "Sample attrition in the Cebu Longitudinal Health and Nutrition Study", (unpublished).

_____ and M. Cinco (2007). "Infant morbidity and school performance in late childhood in Cebu, Philippines", paper presented during the 2007 Annual Meeting of the Population Association of America, New York, 30 March.

Quiza, F. (2006). *Philippine National Health Research System Health Research Agenda of the Visayas 2006-2010*, Bicutan, Tagig, Philippine Council for Health Research and Development.

Vital Horoscope: Longitudinal Data Collection in the Iranian Primary Health Care System

The demographic and health data collected through the use of the vital horoscope in the villages covered by a rural health house provide an extremely rich source of data on a 10-year period of observation.

By Amir H. Mehryar, Mohsen Naghavi, Shirin Ahmad-Nia and Shahla Kazemipour*

Despite their importance, longitudinal methods of data collection are complex, time-consuming and expensive. With longer periods of observation, the problems of population movement and loss to follow-up become important. This is particularly the case in developing countries with high rates of rural-to-urban migration and residential mobility within major urban centres. A more practical

* Amir H. Mehryar, Professor of Population and Behavioral Studies, Institute for Research on Planning and Development, Tehran, e-mail: ahmehryar@gmail.com; Mohsen Naghavi, Senior Epidemiologist, Ministry of Health and Medical Education, Tehran; Shirin Ahmad-Nia, Assistant Professor of Medical Sociology, Department of Sociology, College of Social Sciences, Allameh Tabatabaee University, Tehran; and Shahla Kazemipour, Assistant Professor of Demography, Department of Demography, College of Social Sciences, Tehran University, Tehran.

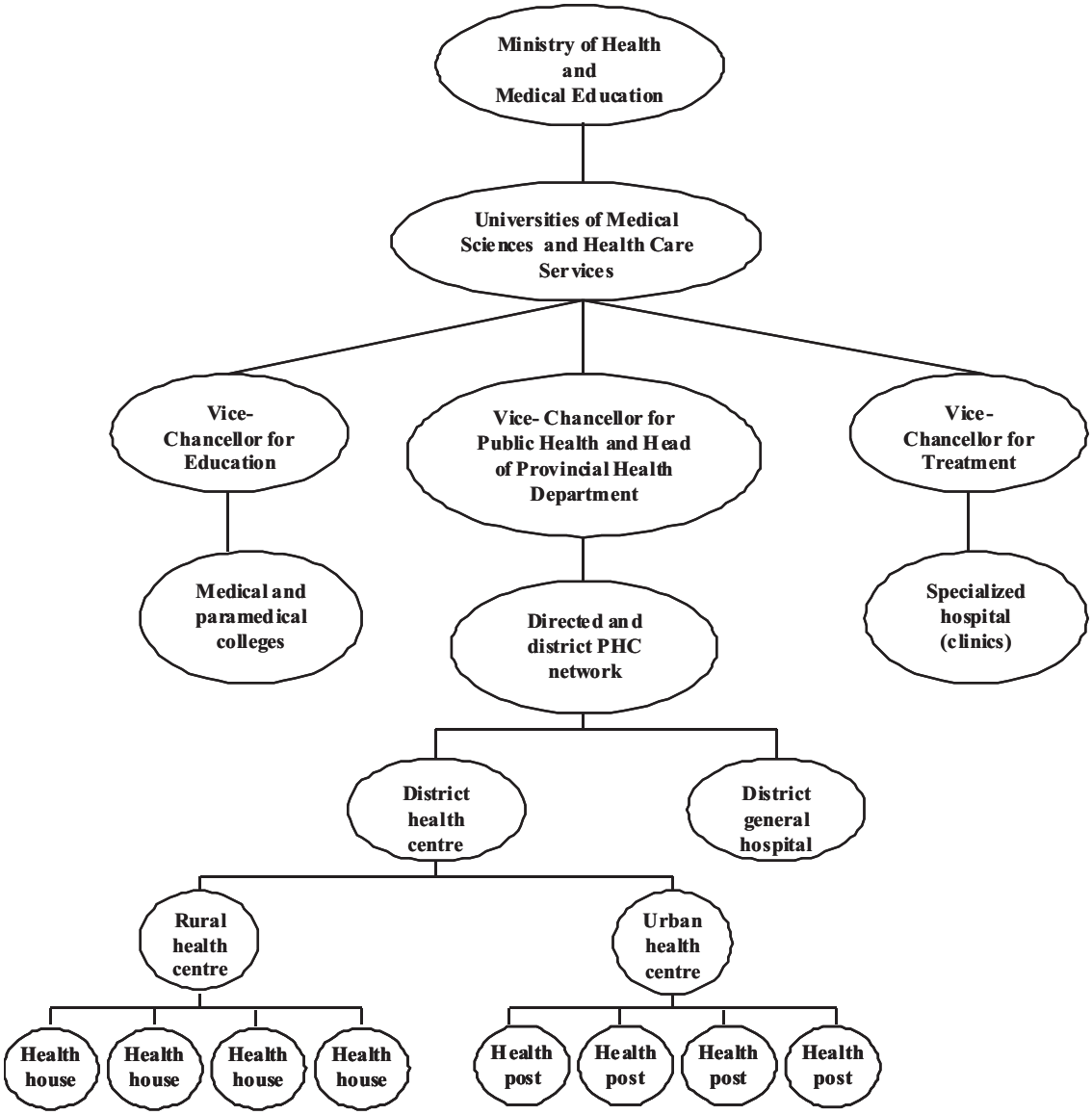
alternative involves repeated data collection on well-defined communities on a regular monthly or annual basis. Combining such data-gathering exercises with a practical programme of intervention or service delivery can justify the expense involved and ensure the viability of the project.

Since its establishment in 1985, the Primary Health Care (PHC) Network of the Islamic Republic of Iran has provided a unique opportunity for longitudinal data collection concerning the total population of a large number of rural communities (villages) on an annual basis (Naghavi and others, 2005). This article strives to describe this PHC-based management information system and show the type of longitudinal information collected by it. To achieve this aim, the article begins with a description of the sophisticated but simple management information system developed to support the PHC Network, the most innovative component of which is the "vital horoscope", which is described at some length in the article. Finally, the trends of changes in basic demographic and health indicators of the rural population of the Islamic Republic of Iran, as revealed by an analysis of vital horoscope data collected between 1993 and 2003, are presented. The combination of cross-sectional and longitudinal data afforded by the vital horoscope is important for obtaining a nuanced and comprehensive picture of the demographic and health status of the population nationally and in different provinces and districts, and for providing valuable information on the reach and effectiveness of the public health services.

The PHC Network and health information system

The foundations of the PHC Network, as the backbone of the public health system of the Islamic Republic of Iran, were laid in 1984 when a law passed by the Islamic Consultative Assembly (parliament) drastically changed the structure and organization of the Ministry of Health. According to that law, all medical and paramedical universities and training centres in the country were separated from the Ministry of Higher Education and put under the direct control of the Ministry of Health (renamed the Ministry of Health and Medical Education. As indicated in figure 1, that Ministry is the national authority responsible for training medical personnel, gathering health information and providing health-care services. Its human resources service delivery and monitoring/evaluation functions are implemented by the provincial Universities of Medical Sciences and Public Health Care Services (UMSHCS) that have health-care responsibilities in their respective areas. Currently, there are 40 publicly funded UMSHCS that are run by the Ministry covering the country's 30 provinces, with seven provinces (Esfahan, Fars, Kerman, Mazandaran, Semnan, Sistan-Baluchestan and Tehran) having two or more UMSHCS run by the Ministry.

Figure 1. Organizational structure of the Iranian health system



Each province is divided into several subprovinces or districts that include a few urban centres plus a large number of villages. With regard to the administration of health services, the district health department is the autonomously functioning administrative unit at the most basic level. The lowest and most essential parts of the district health network are the rural health houses and urban health posts, both of which are designed to deliver the variety of services envisaged by the PHC concept at

the grass-roots rural and urban levels, respectively. In addition to the rural health houses, the rural section of the PHC system also includes rural health centres and maternity facilities affiliated with the rural health houses.

In particular, each village-based rural health centre, headed by a physician and staffed by paramedical and administrative staff, supervises the rural health house in its own village and usually a few others in neighbouring villages, covering a total of about 7,500 people. The chief responsibilities of a rural health centre are to support and supervise rural health house activities, accept referred cases and maintain proper contact with higher levels of the health system, particularly the district general hospital. Located in individual villages, each rural health house is normally the closest, and usually the only, health facility available in the area, serving an average target population of 1,500 people living there and in one or more satellite villages (no more than one hour's walk away from the centre). Each rural health house is staffed by one or more female health worker and one male health worker known as a *behvarz* (a neologism meaning "health provider") chosen from the same village where she/he is to be stationed in the future and trained at a *behvarz* training centre located in each district. The student health providers (*behvarzan*) receive free training and financial support throughout the two-year period of their training in return for serving at the rural health house for a minimum of four years after completing their training. The female *behvarz* is generally responsible for indoor tasks, that is, receiving clients, dispensing routine health care and simple curative care to those under her responsibility, including immunizing clients, recording data and providing patients with medications, while the male *behvarz* is concerned with outdoor activities, that is, following up cases with communicable diseases, preparing case findings, administering immunizations, overseeing environmental health and providing routine health care in satellite villages.

The *behvarzan* in the rural health house, as well as in the urban health house, are an important component of the health information system that was developed to collect detailed information on the community and the population served by the PHC Network and to evaluate the effectiveness and impact of its interventions. At the local level, it is the *behvarz* who is entrusted with the task of collecting, recording and storing health information to feed into the health information system according to various scheduled activities, including the taking of an annual census of the population of the catchment area at the beginning of each year (late March-early April) and keeping individual files on each household entrusted to his or her care. Those files contain the name, sex, age, literacy status and level of education of each household member, the mother's condition during any previous

pregnancies, the health status of the children, the types of health care delivered to each child, the major diseases that have afflicted household members, and the births and deaths that have occurred during the preceding year. Additional data on the dwelling unit, the number of rooms and the availability of electricity, piped water, toilets, heating and the like are also collected. The female *behvarz* is further required to keep a record of all her daily activities either in the family files or in a specifically designed logbook. There are separate logbooks for specific activities, such as family planning, medications dispensed and immunizations. In order to summarize and graphically present the results of the data collected by each *behvarz,* an innovative device called the "vital horoscope" has been designed.

Vital horoscope

The vital horoscope which is printed on a 50 x 70 cm sheet of paper, is designed to display an up-to-the-hour account of the vital events recorded by the *behvarz* and the services provided by the health house during the year. The chart owes its name to the concentrically coloured circles in the centre, which resemble an ancient astrological horoscope. The innermost circle allocates space for recording the year covered by the horoscope. From the centre outward, the remaining four circles represent live births, deaths among infants, deaths among children aged 1-5 years and deaths of persons older than 5 years, respectively. Each circle is divided into 12 segments, representing the 12 months of the year. Within each of these sections small squares and circles are used to represent data from the main village and the satellite villages covered by the health house. The vital horoscope also includes seven tables printed around the main circle. The first table contains the urban or rural identity of the health house, the number of households it serves and the number of households using iodized salt. The second table presents a detailed summary of the annual census taken at the beginning of each year and shows the age and sex structure of the population of the main and satellite villages served by the health house. In this table, the number of currently married women aged 15-49 years is given separately. The third table summarizes the number of births by outcome, such as live births and stillbirths, sex and weight of the baby, age of the mother, place of delivery (hospital/delivery centre or home) and persons helping with the delivery (trained or untrained). The fourth table records the number of deaths associated with the pregnancy or delivery, specifying the four major causes of death, and the age of the mother. The fifth table summarizes information on the coverage of the family planning programme, such as the number of couples using different methods of contraception, at the end of each three-month period. The sixth table summarizes the number of deaths by sex and age. The seventh table gives more detailed information on deaths among children

under five years by age and eight major causes of death. All the data included in these tables are listed separately for the main and satellite villages. Comparing the figures recorded in the circle with those on the adjacent tables provides the opportunity to maintain quality control rapidly at the lowest administrative level.

When it was first introduced in 1988, the main purpose of the vital horoscope was to help the female *behvarz* summarize her activities and to acquire an up-to-date picture of the health and demographic characteristics of the population under her care. The device could also be used to help inspectors and administrators at higher levels of the PHC programme to monitor the activities of individual *behvarz* as well of the rural health house and rural health centre, and to assess the impact of the entire programme on the health status of the rural population. Each vital horoscope is designed to cover one calendar year. Throughout the year, it remains pinned on the wall of the rural health house; at the end of each year, when a new vital horoscope sheet is issued, a copy of the past year's vital horoscope is sent to the district health department for analysis and comparison with those of other facilities under its supervision.

Until recently, only the summary results of vital horoscopes analysed by different health districts were sent to the central office of the PHC Network in Tehran. Since 1999 special software has been developed for analysing the vital horoscopes; it is used by the central office and the provincial and district-level PHC departments. Apart from occasional oral reports by interested *behvarzan*, there is no provision for sharing the results of the vital horoscope with the rural community that it covers. Until recently, no effort had been made to share the data collected by the rural health house with the academic and research community outside the Ministry. The recently published monograph by Naghavi and others (2005) is the most detailed analysis of the vast amount of data derived from vital horoscopes completed between 1993 and 2003. It reveals impressive changes in various indicators of the demographic composition and health status of the population covered by the rural health houses. The enormous amount of data gathered through vital horoscopes make possible many types of analysis, but in this article only trends over time and comparison across different locations (28 provinces and 280 districts) are examined.

Analysis of vital horoscope data for the period 1993-2003

Changes in the number of rural health houses and population coverage

Since the introduction of the PHC Network in 1984, the number of rural health houses and *behvarzan* has been growing steadily. By 2003 there were some 16,000 rural health houses serving a population of over 20 million, which

represents more than 85 per cent of the rural population of the Islamic Republic of Iran. With the increase in the number of rural health houses, the quality of services offered by the *behvarzan* has also improved. This is reflected in a decrease in the mean number of the population covered, from 1,519 to 1,255 between 1993 and 2003 (see table 1), and a fall in the ratio of people living in satellite villages of the population covered by each *behvarz*. Thus, the *behvarzan* have become increasingly free of the burden of travelling to satellite villages. The share of satellite villages of the population covered by each rural health house in 2003 varied from 51 per cent in Kohgiluyeh-Boyerahmad Province to 9 per cent in East Azarbayjan Province.

Table 1. Changes in the number of rural health house and average size of population served by each house between 1993 and 2003

1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Number of active rural health houses										
11 029	12 061	12 626	13 753	13 928	14 879	15 218	15 406	15 916	15 947	16 065
Population covered										
1 519	1 464	1 473	1 417	1 449	1 374	1 368	1 339	1 297	1 292	1 255

Demographic changes

Data from vital horoscopes confirm the results of other surveys, indicating that the changes in the age structure and fertility rates of the Iranian population since the late 1980s have been dramatic (see table 2). This decline is noticeable in the relative share of children under 12 months of age (from 2.71 per cent in 1993 to 1.63 per cent in 2003) as well as the share of children aged 0-14 years (from 44.7 per cent in 1993 to 31.6 per cent in 2003). In contrast, the share of the economically active population aged 15-64 years has risen (from 51 per cent to 61 per cent) while the share of the elderly aged 65 years or older in the rural population covered by rural health houses has also increased (from 4.2 per cent in 1993 to 6.3 per cent in 2003). As a result, the total dependency ratio has dropped from 96.2 in 1993 to 61.4 in 2003. Significant changes are also noticeable in the proportion of women of reproductive age (15-49 years) as well as the proportion of women married in the age groups 15-19, 20-24, 25-29 and 30-34 (see table 3).

Table 2. Changes in age and sex composition of rural population covered by the rural health houses between 1993 and 2003

Index	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Percentage aged 0-14	44.7	43.9	42.7	41.5	40.2	39.0	37.6	36.1	34.4	32.9	31.6
Percentage aged 15-64	51.1	51.8	53.8	53.8	55.8	55.8	57.0	58.4	60.8	60.9	61.1
Percentage aged 65+	4.2	4.3	4.5	4.7	5.0	5.2	5.4	5.5	5.8	6.2	6.3
Sex ratio at birth (of live births)	105	105	105	105	104	105	106	105	106	104	106
Sex ratio at birth (of stillborn)	133	130	131	133	130	128	130	129	120	124	123
Sex ratio of neonatal deaths	122.7	132.0	120.1	122.9	126.2	126.3	131.8	129.0	134.1	130.2	128.0
Dependency ratio	96.2	93.8	90.1	86.2	82.9	79.7	75.7	71.7	67.8	64.4	61.4

Table 3. Changes in age-specific marriage rates of women in rural areas covered by the rural health houses between 1993 and 2003

Index Percentage of married women aged:	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
15-19	19.8	18.5	17.0	16.0	15.3	14.4	14.0	15.6	16.0	16.0	16.1
20-24	57.4	56.4	55.1	53.6	52.1	50.4	49.2	49.4	48.8	48.5	48.0
25-29	79.7	78.8	78.8	77.9	76.8	75.7	74.8	73.9	73.0	71.9	70.7
30-34	88.5	88.4	88.0	87.8	86.9	86.8	86.0	85.3	84.7	83.7	83.1
10-49	**63.0**	**62.4**	**61.4**	**60.6**	**59.8**	**59.2**	**58.7**	**59.0**	**58.7**	**58.3**	**58.0**

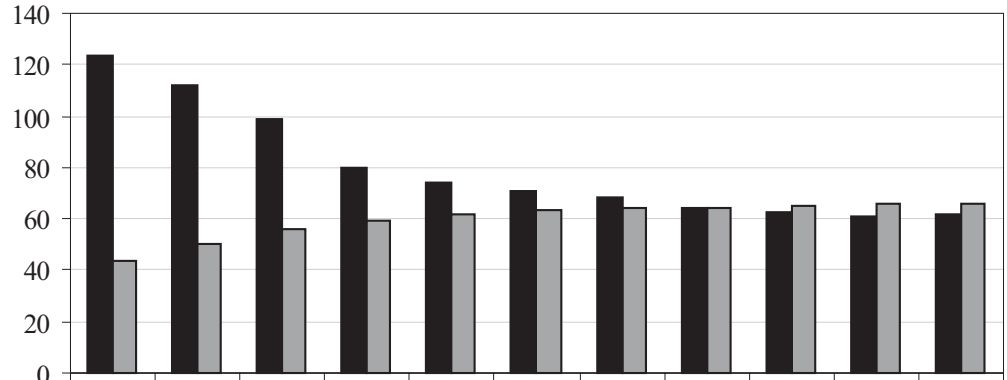
Changes in fertility rates

As a result of these changes, the fertility rates of women in rural areas has fallen drastically (see table 4). This is clearly reflected in the drop in the total fertility rate (from 3.9 live births per woman to 2.0), the general fertility rate (from 124 to 62 live births per 1,000 women of childbearing age) and the crude birth rate (from 24.1 to 16.3 live births per 1,000 populations) between 1993 and 2003. That these changes were mainly due to the widespread use of modern contraceptives promoted and provided by the *behvarz* is clearly indicated by the fact that the prevalence of modern contraceptive use has risen from 43.7 per cent to 66.2 per cent among women aged 15-49 years between 1993 and 2003 (see figure 2).

Table 4. Changes in measures of fertility, mortality and population growth of villages covered by the rural health houses in the Islamic Republic of Iran

Index	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Total fertility rate (live births per woman)	3.9	3.4	3.0	2.6	2.4	2.3	2.2	2.0	2.0	1.9	2.0
General fertility rate (live births per 1,000 women 15-19 years old)	123.8	112.1	98.7	80.0	74.3	71.2	68.6	64.5	63.0	61.1	61.9
Crude birth rate (live births per 1,000 population)	24.1	22.2	20.0	17.9	17.0	16.6	16.4	15.8	15.9	15.8	16.3
Crude death rate (deaths per 1,000 population)	4.5	4.3	4.7	4.2	4.4	4.2	4.4	4.4	4.5	4.8	5.1
Rate of natural increase (percentage)	1.96	1.79	1.53	1.37	1.26	1.24	1.20	1.14	1.14	1.00	1.12
Contraceptive prevalence rate	43.7	50.5	56.1	59.5	61.6	63.3	64.5	64.4	65.1	65.5	66.2

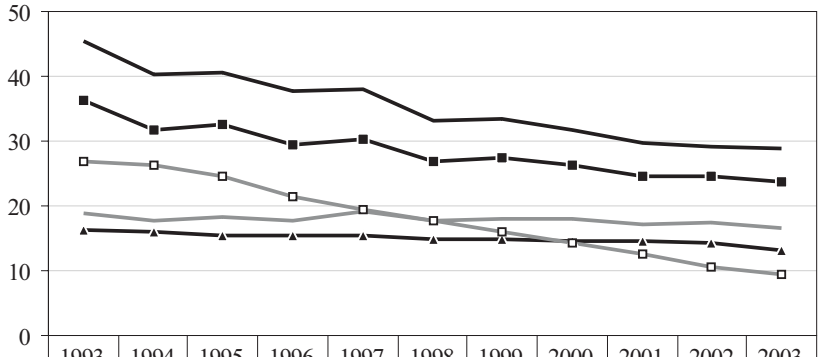
Figure 2. General fertility and contraceptive prevalence rates of women in rural areas of the Islamic Republic of Iran, 1993-2003



	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
GFR (per 1,000)	123.8	112.1	98.7	80.0	74.3	71.2	68.6	64.5	63.0	61.1	61.9
CPR (per cent)	43.7	50.5	56.1	59.5	61.6	63.3	64.5	64.4	65.1	65.5	66.2

Note: GFR stands for general fertility rate
CPR stands for contraceptive prevalence rate

Figure 3. Changes in stillbirths, neonatal mortality, infant mortality, under-five mortality rates and proportion of deliveries by untrained personnel, 1993-2003



	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Stillbirths	16.4	16.0	15.5	15.4	15.3	14.8	14.9	14.7	14.5	14.3	13.1
NNMR	19.0	17.6	18.2	17.8	19.1	17.8	18.1	18.0	17.2	17.4	16.6
IMR	36.3	31.7	32.5	29.5	30.2	26.8	27.4	26.3	24.6	24.5	23.7
U5MR	45.3	40.2	40.6	37.8	37.9	33.2	33.5	31.6	29.7	29.1	29.0
Percentage of DUP	27.0	26.4	24.7	21.3	19.3	17.6	16.1	14.2	12.7	10.7	9.3

Note: NNMR stands for neonatal mortality rate
IMR stands for infant mortality rate
U5MR stands for under-five mortality rate
DUP stands for deliveries by untrained personnel

Variations

Changes in basic health indicators

Noticeable improvements in infant and child care were also recorded. As figure 3 shows, between 1993 and 2003 the number of deliveries by unskilled personnel had decreased, as had the number of stillbirths, dropping from 16.4 stillbirths per 1,000 births to 13.1, and the neonatal mortality rate, falling from 19 deaths of infants under 28 days of age per 1,000 live births in 1993 to 16.6 in 2003. Infant mortality rates dropped from 36.3 deaths of infants under 1 year of age per 1,000 live births in 1993 to 23.7 in 2003.

Provincial variations

Despite the drastic changes in these various demographic and health indicators, it should be noted that provincial variations remain with respect to almost all these variables. This is not unexpected in view of the diversity in ethno-cultural background and in the level of development of the provinces in the Islamic Republic of Iran. For instance, in 2003 the share of children under 15 years of age in the rural population covered by rural health houses varied from 42 per cent in Sistan-Baluchestan Province to 25 per cent in Mazandaran Province. Similarly, the share of the elderly of the population of different provinces varied from 11 per cent in Qom and Markazi provinces to 4 per cent in Sistan-Baluchestan, Golestan and Khuzastan provinces. Consequntly, the dependency ratios of provinces vary between 86 in Sistan-Baluchestan to 48 in Mazandaran (see table 5).

Table 5. Provincial variations in age and sex composition of the rural population covered by the rural health houses in 2003

Province	Percentage of population aged:					Dependency ratio	Sex ratio
	<1 year	<5 years	0-14 years	15-64 years	65+ years		
East Azarbaijan	1.63	7.8	30	63	7	57	106
West Azarbaijan	1.86	9.0	34	60	6	65	106
Ardebil	1.50	7.5	32	62	6	62	105
Esfehan	1.33	6.7	27	65	8	53	106
Ilam	1.37	6.1	29	66	5	52	100
Bushehr	1.80	8.4	32	63	5	60	104
Tehran	1.59	7.9	29	65	6	52	106
Chaharmahal-Bakhtiary	1.74	8.7	35	60	5	67	106
Khorasan	1.71	8.7	33	60	7	67	105

.../

Table 5 *(Continued)*

Province	Percentage of population aged:					Dependency ratio	Sex ratio
	<1 year	<5 years	0-14 years	15-64 years	65+ years		
Khuzestan	2.04	10.1	37	59	4	71	105
Zanjan	1.57	7.6	32	62	6	62	105
Semnan	1.43	7.1	27	63	10	60	112
Sistan-Baluchestan	2.67	13.0	42	54	4	86	105
Fars	1.54	7.5	31	63	6	57	106
Qazvin	1.55	7.2	29	64	7	58	102
Qom	1.28	6.8	26	63	11	60	111
Kordestan	1.54	7.6	34	60	6	65	105
Kerman	1.68	8.4	33	61	6	64	107
Kermanshah	1.38	7.0	31	63	6	58	108
Kohgiluye-Boyerahmad	1.73	8.6	35	60	5	67	110
Golestan	1.63	8.2	33	63	4	60	105
Gilan	1.20	6.2	25	67	8	51	106
Lorestan	1.59	7.9	34	61	5	65	107
Mazandaran	1.30	6.5	25	68	7	48	107
Markazi	1.23	5.9	26	63	11	58	107
Hormozgan	2.03	10.2	37	57	6	76	105
Hamedan	1.50	7.3	31	62	7	60	102
Yazd	1.55	7.4	27	63	10	60	95
Iran (Islamic Republic of)	**1.63**	**8.1**	**32**	**62**	**6**	**61**	**105**

The total fertility rates of the provinces varied from 3.6 in Sistan-Baluchestan to 1.3 in Gilan, with the average for all provinces being 1.9. Similar variations are noted in the prevalence of modern contraceptive use, which in 2003 varied from 80 per cent in Zanjan to 57 per cent in Gilan, with only two provinces falling under 50 per cent (Sistan-Baluchestan at 45 per cent and Hormozgan at 42), with the overall mean for all provinces being 66 per cent (see table 6).

Table 6. Provincial variations in various fertility indicators of women in rural areas covered by rural health house services in 2003

Province	Crude birth rate (per 1,000)	Total fertility rate (per woman)	General fertility rate (per 1,000)	Modern contraceptive prevalence (per cent)	Percentage married of women aged 15-49
East Azarbaijan	16	1.9	61	70	59
West Azarbaijan	18	2.2	70	74	59
Ardebil	15	1.8	58	75	55

.../

Table 6 *(Continued)*

Province	Crude birth rate (per 1,000)	Total fertility rate (per woman)	General fertility rate (per 1,000)	Modern contraceptive prevalence (per cent)	Percentage married of women aged 15-49
Esfehan	13	1.5	50	71	62
Ilam	14	1.7	50	78	45
Bushehr	18	2.2	67	62	52
Tehran	13	1.4	47	63	63
Chaharmahal-Bakhtiary	19	2.3	74	70	59
Khorasan	17	2.2	67	60	60
Khuzestan	21	2.5	81	67	55
Zanjan	16	1.8	59	80	58
Semnan	14	1.6	53	66	61
Sistan-Baluchestan	26	3.6	109	45	62
Fars	16	1.8	56	66	54
Qazvin	15	1.7	57	70	58
Qom	12	1.5	49	57	64
Kordestan	16	2.0	63	74	61
Kerman	16	1.9	59	59	52
Kermanshah	15	1.8	55	76	53
Kohgiluye-Boyerahmad	18	2.2	68	69	54
Golestan	17	2.0	63	70	57
Gilan	12	1.3	43	63	58
Lorestan	17	2.0	65	75	55
Mazandaran	13	1.3	44	66	62
Markazi	13	1.5	48	70	54
Hormozgan	20	2.6	78	42	58
Hamedan	15	1.8	59	73	60
Yazd	14	1.9	56	67	61
Iran (Islamic Republic of)	**16**	**1.9**	**61**	**66**	**58**

The vital horoscope data are also of help in recording the impact of natural disasters and other exceptional events. In 2003, as table 7 shows, provincial crude death rates (CDR) varied from 8.1 deaths per 1,000 population in Kerman Province to 3.8 in Kohgiluye-Boyerahmad Province, with the average for all provinces being 5.1. The rise in CDR experienced by Kerman Province between 2002 (4.4) and 2003 (8.1) was caused by the Bam earthquake, which resulted in more than 28,000 deaths. In 2002, provincial differences in CDR ranged from 6.4 in Semnan to 3.8 in Tehran, with the overall rate being 4.8. Kerman Province ranked twentieth in CDR (4.4).

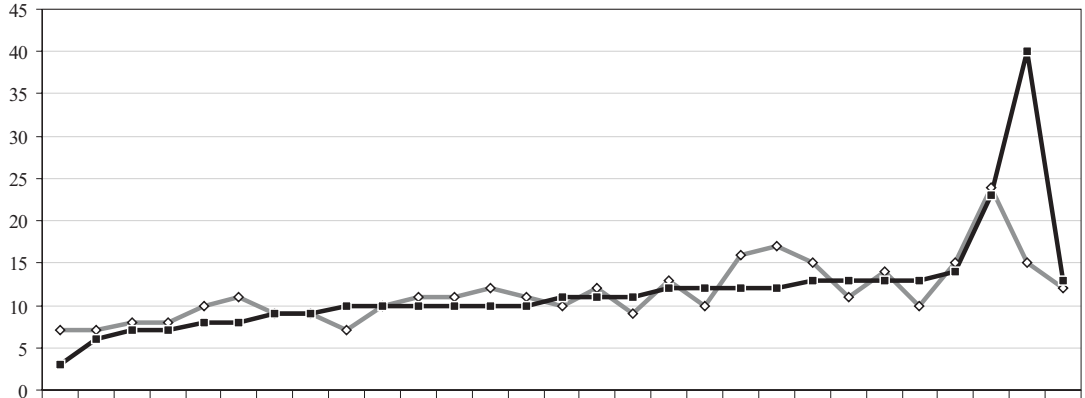
Table 7. Provincial variations in different mortality rates of villages covered by the rural health house services in 2003

Province	Crude death rate (per 1,000)	Neonatal mortality rate (per 1,000)	Infant mortality rate (per 1,000)	Under-5 mortality rate (per 1,000)	Mortality rates of children (per cent)
East Azarbaijan	5.1	16	23	27	11
West Azarbaijan	4.6	17	24	30	13
Ardebil	4.8	17	23	28	11
Esfehan	5.1	17	22	26	9
Ilam	4.0	12	17	22	10
Bushehr	5.0	20	26	29	9
Tehran	3.9	7	11	14	6
Chaharmahal-Bakhtiary	4.3	16	22	26	10
Khorasan	5.5	18	27	32	14
Khuzestan	3.9	16	24	29	13
Zanjan	5.1	16	21	24	8
Semnan	7.3	14	19	24	10
Sistan-Baluchestan	4.9	18	33	41	23
Fars	4.5	15	22	27	12
Qazvin	5.3	17	22	25	8
Qom	6.3	8	9	11	3
Kordestan	4.8	23	29	33	10
Kerman	8.1	16	29	55	40
Kermanshah	4.7	17	25	30	13
Kohgiluye-Boyerahmad	3.8	18	25	30	12
Golestan	4.7	18	26	31	12
Gilan	6.2	17	21	24	7
Lorestan	4.6	24	31	37	13
Mazandaran	5.2	11	15	18	7
Markazi	6.5	13	19	24	11
Hormozgan	4.5	21	27	32	12
Hamedan	5.3	18	25	28	10
Yazd	5.9	13	18	22	10
Iran (Islamic Republic of)	**5.1**	**17**	**24**	**30**	**13**

The impact of the earthquake on children in Kerman can also be observed in child and under-five mortality. In 2003, provincial variations in the mortality rates of children aged 1-59 months ranged from 40 deaths per 1,000 live births in Kerman Province to 3 in Qom, with the mean being 13. In 2002, Kerman Province had a child mortality rate of 15 (see figure 4). Provincial variations in the under-five mortality rates in 2002 ranged from 45 per 1,000 live births in

Sistan-Baluchestan Province to 15 in Tehran Province, with the overall mean being 29. For Kerman Province, the under-five mortality rate was 32, which was much lower than the 2003 figure of 55 (see figure 5).

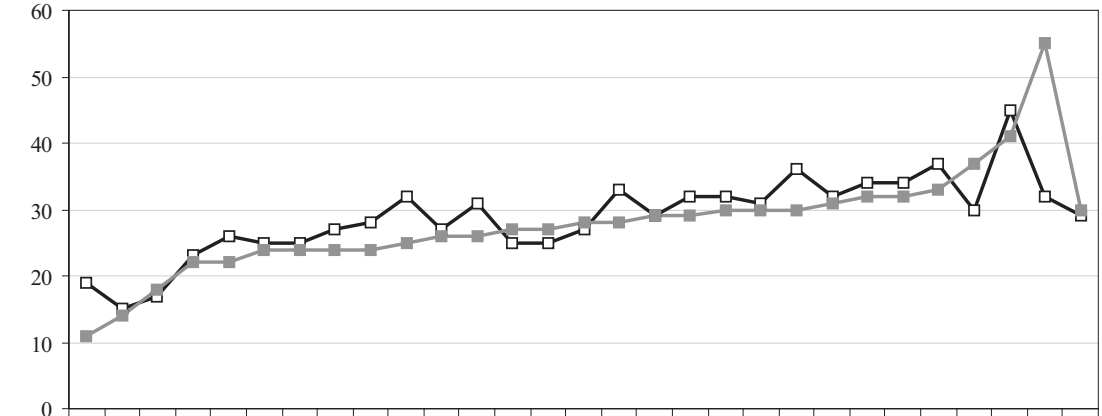
Figure 4. Provincial variations in child mortality rates (children aged 1-59 months) of the rural population covered by rural health houses in 2002 and 2003



	Qom	Tehran	Mazandar	Gilan	Zanjan	Qazvin	Esfehan	Bushehr	Yazd	Hamedan	Semnan	Ilam	Chaharm	Kurdesta	Markazi	Ardebil	E. Azarb	Golestan	Fars	Hormozg	Khgiluy	Khuzesta	Lorestan	W. Azarb	Kermans	Khorasan	Sistan&B	Kerman	Iran
CMR 2002	7	7	8	8	10	11	9	9	7	10	11	11	12	11	10	12	9	13	10	16	17	15	11	14	10	15	24	15	12
CMR 2003	3	6	7	7	8	8	9	9	10	10	10	10	10	10	11	11	11	12	12	12	12	13	13	13	13	14	23	40	13

Note: CRM stands for child mortality rate

Figure 5. Provincial variations in the under-five mortality rates of the rural population covered by rural health houses in 2002 and 2003



	Qom	Tehran	Mazan	Yazd	Ilam	Gilan	Markazi	Zanjan	Semnan	Qazvin	Esfeha	Chahar	E. Azar	Fars	Hameda	Ardebil	Bushehr	Khuzest	Kerman	W. Azar	Khgilu	Golesta	Khoras	Hormoz	Kurdest	Loresta	Sistan	Kerman	Iran
U5MR 2002	19	15	17	23	26	25	25	27	28	32	27	31	25	25	27	33	29	32	32	31	36	32	34	34	37	30	45	32	29
U5MR 2003	11	14	18	22	22	24	24	24	24	25	26	26	27	27	28	28	29	29	30	30	30	31	32	32	33	37	41	55	30

Note: U5MR stands for under-five mortality rate

District-level variations

An analysis of the vital horoscope data also indicates significant variations among the 280 subprovinces or districts in terms of various demographic and health indicators. These differences are clearly reflected in table 8, which summarizes descriptive statistics of different variables for the three-year period 2000-2002. Neonatal mortality rates for that three-year period varied from 3.6 to 41.5, with the median being 17.5 deaths of infants less than 28 days old per 1,000 live births. Infant mortality rates for the three-year period varied from 6.0 to 70.2, with the median being 24.3 deaths of infants under 1 year of age per 1,000 live births. The infant mortality rate was above 30 such deaths in 59 districts, above 40 in 8 districts and above 50 in only two districts. The under-five mortality rates for the three-year period varied from 9.6 to 83.5, with the median being 29.3 per 1,000 live births. That rate was above 30 in 121 districts; it exceeded 40 in 31 districts and was above 50 in only 5 districts.

Table 8. Summary of descriptive statistics of selected variables for 2000-2002

	Number	Mean	Standard deviation	Variance	Range (Percentage)
Percentage share of population aged under 1 year	281	1.5993	0.36907	0.136	0.8-3.1
Percentage share of population aged under 5 years	279	8.1269	1.84554	3.406	3.7-15.1
Percentage share of population aged 0-14 years	279	33.6079	4.85772	23.597	19-4-45.7
Percentage share of population aged 65+ years	277	6.2783	2.46142	6.059	2.6-18.3
Percentage share of population aged 60-69 years	279	6.3792	2.08651	4.354	3.2-15.3
Percentage share of population aged under 70-79 years	279	3.1272	1.31112	1.719	0.2-10.3
Percentage share of population aged 80+ years	279	0.7774	0.50375	0.254	0.1-3.1
Percentage share of population aged 15-64 years	279	59.8591	4.55386	20.738	45.9-86.0
Percentage married of women aged 15-49 years	279	57.8975	5.47406	29.965	43.4-73.5

.../

Table 8 *(Continued)*

	Number	Mean	Standard deviation	Variance	Range (Percentage)
Percentage married of women aged 15-20 years	279	15.1552	5.66954	32.144	2.7-30.0
Percentage married of women aged <15 years	279	0.5921	0.53665	0.288	0.0-5.3
Crude birth rates (per 1,000)	279	15.3470	3.41077	11.633	6.5-26.5
General fertility rate of women aged 15-49 years	279	60.8473	15.93091	253.794	28.9-125.3
Total fertility rates of women aged 15-49 years	279	1.9405	0.54645	0.299	1.0-4.5
Age-specific fertility rate (15-19 years)	279	28.4674	11.47125	131.590	5.9-74.3
Percentage of pregnancies identified as high risk due to the age of the mother (<20 years and >35 years).	279	21.5391	3.53234	12.477	12.5-32.6
Modern contraceptive prevalence rate (%)	279	64.3584	10.12408	102.497	21.0-80.0
Percentage of deliveries by untrained personnel	279	8.9548	10.29671	106.022	0.0-57.4
Percentage of deliveries at home	279	17.0326	21.38042	457.122	0.0-89.8
Dependency ratio	279	67.3061	10.18778	103.791	46.9-103.5
Crude death rate (both sexes) (per 1,000 population)	279	4.7459	1.11027	1.233	2.0-9.5
Neonatal mortality rate (per 1,000 live births)	279	17.7240	5.54720	30.771	3.6-41.5
Infant mortality rate (per 1,000 live births)	279	24.8545	7.78498	60.606	6.0-70.2
Under-five mortality rate (per 1,000 live births)	279	29.9695	9.12864	83.332	9.6-83.5
Child mortality rate (mortality of children aged 1-59 months per 1,000 live births)	279	12.2527	5.59582	31.313	2.7-42.7
Age-specific death rate (people aged 15-50 years) (per 1,000)	279	13.4054	2.85400	8.145	5.9-23.5
Maternal mortality ratio (per 100,000 deliveries)	279	31.7097	40.64188	1 651.763	0.0-174

Maternal mortality ratios for the three-year period 2000-2002 varied from zero to 174 deaths of women per 100,000 live births, with the median being 16.3. This ratio was zero in 134 districts, below 30 in 37 districts, between 30 and 50 in 34 districts, between 50 and 80 in 36 districts, between 80 and 100 in 17 districts, and above 100 in 23 districts, rising above 150 in only 6 of those 23 districts. The prevalence rates of modern contraceptive use varied from 80 per cent to 21 per cent, with the median being 65.7 per cent. That rate was above 70 per cent in 66 of the 281 districts, between 60 per cent and 69.9 per cent in 110 districts, below 50 per cent in 23 districts and below 40 per cent in only 9 of the 281 districts.

Discussion and conclusions

The demographic and health data collected through the use of the vital horoscope in the villages covered by a rural health house provide an extremely rich source of data on a 10-year period of observation. This report offers only a glimpse of the data analysed so far. The data indicate remarkable changes in most of the demographic and health indicators of the population covered by the rural health houses. On the whole, there is strong evidence of a continuing drop in fertility rates, a decrease in the share of children aged 0-14 years in the total population, a considerable rise in the proportion of the population aged 15-64 years and a slight rise in the share of the elderly aged 65 years and older. The concomitant rise in modern contraceptive use rates confirms the major role played by family planning in the fertility decline and the age-structure changes being experienced by the rural population of the Islamic Republic of Iran.

These findings are entirely in line with the findings of the huge demographic health survey-type assessment conducted in late 2000 (Islamic Republic of Iran, 2003; Mehryar and others, 2003), which indicated that the crude birth rates of Iranian women had dropped to 16.3 births per 1,000 population, that is, 15.2 and 18.4 in urban and rural areas, respectively. According to that large-scale survey, by late 2000 almost 74 per cent of married couples (77.4 per cent of urban couples and 67.2 per cent of rural couples) were using contraceptives. The proportion of rural women using modern contraceptives (57.3 per cent) was higher than that of urban women (55.2 per cent), with a larger proportion of rural women using the traditional method of withdrawal (22.2 per cent vs. 10.0 per cent). The national census conducted in November 2006 offers further support of the accuracy of the vital horoscope in that it indicates a drastic fall in the share of children aged 0-14 years in the total population, from 39.57 to 25.08 per cent between 1996 and 2006. Equally sharp declines are noted in both urban areas (from 37.55 to 23.71 per cent) and rural areas (from 42.57 to 28.04). The crude birth rates of the urban and rural

populations derived from the 2006 census are 15.45 and 17.80 per 1,000 population, respectively (Mehryar, 2007). The difference between the 2006 census figure for rural areas (a CBR of 17.8) and the figure obtained by the vital horoscope data for the year 2003 (a CBR of 16.3) is partly due to time differences and to the fact that the census data refer to the total population of rural areas while the vital horoscope data are based on births reported by women covered by the rural health house services only.

Provincial and district-level variations in various indices derived from the vital horoscope data may be interpreted as a reflection of this validity. Iranian provinces vary considerably in terms of level of development, cultural traditions and ethnic composition of their populations. Human development indices based on the relative standing of different provinces with respect to income, health and education vary from 0.842 in Tehran Province to 0.545 in Sistan-Baluchestan, the national average being 0.76 (Plan and Budget Organization of Iran and United Nations Development Programme in Iran, 1999). In general, the rural health houses in the country's northern provinces (Gilan and Mazandaran) and central provinces (Tehran, Qom, Esfahan, Markazi, Semnan and Yazd) with higher levels of development have reported lower rates of fertility and infant mortality but higher rates of contraceptive prevalence. The reverse is seen in the case of southern provinces (Hormozgan and Kohgiluye-Boyerahmad), south-eastern provinces (Sistan-Baluchestan and Kerman) and north-western provinces (West Azarbaijan, Kurdistan, Kermanshah and Ilam). The variations are entirely consistent with other evidence regarding the level of development of the provinces. A similar pattern is noted with regard to district-level variations. On the whole, despite differences in population size and urbanization rates, districts within each province are very similar to the overall average of the province with respect to most measures of health and development.

It is worth noting that the impact of the rural health houses would seem to vary by the nature of the health measures used. For example, infant mortality rates would appear to have dropped more markedly and consistently between 1993 and 2003 than neonatal mortality and stillbirth rates. The drop in the stillbirth rate is not only less striking but also marked by variations over time. The kind of skills and services provided by the *behvarz* would seem to be more effective in dealing with infant mortality than neonatal or perinatal mortality. In the case of the crude birth rate and the contraceptive prevalence rate, it would seem that a plateau had been reached in 1997, since which time there has been little change.

Acknowledgements

The authors would like to express their gratitude to Dr. Muhammad Ali Mirzaee, Chairman, Afzalipour Foundation for the Promotion of Scientific Research, Kerman, Islamic Republic of Iran, for providing partial support for the project on which this article is based.

References

Islamic Republic of Iran (2003). *Simaye Jamiyat va Salamat dar Jumhuri Islami Iran, 1379,* (A Profile of Population and Health in the Islamic Republic of Iran, 2000), Tehran, Ministry of Health and Medical Education (in Persian).

Mehryar, A.H. and others (2003). *Demographic and Health Survey of Iran (2000): A Summary of the Main Findings,* Tehran, Center for Population Studies and Research, Working Paper Series in English, No. 6.

Mehryar, A.H. (2007). *Highlights of Iranian Census 2006, Tehran*, Center for Population Studies and Research, Working Paper Series in English, No. 21.

Naghavi, M. and others (2005). *Degargoni Simaye Salamat dar Roosta Neshinan Iran, 1372-1382* (Changes in the Health Profile of the Rural Population of Iran, 1993-2003), Tehran, Ministry of Medical Health and Education (in Persian).

Plan and Budget Organization of Iran and the United Nations Development Programme in Iran (1999). *Human Development Report of the Islamic Republic of Iran, 1999,* Tehran, UNDP.

The Effects of Intergenerational Support on the Changes in Cognitive Functioning of the Rural Elderly in China

Rapid economic growth, accelerating migration patterns and changing family preferences in China have jeopardized the ability of older parents to rely on their adult children for needed support and this situation is compromising the well-being of the elderly.

By Ping Wang and Shuzhuo Li*

A major issue in all societies, especially those that are ageing, is the extent to which the elderly can live independently. Cognitive functioning is a crucial factor in the degree of independence of the elderly, while cognitive impairment can predict the loss of functions and even death. There is a close and direct relationship between social support and cognitive functioning, and social support-preventable

* Ping Wang, School of Management, Xi'an Jiaotong University, Xi'an University of Science and Technology, e-mail: wplqym@163.com, and Shuzhuo Li, Institute for Population and Development Studies, School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an, Shaanxi, China.

cognitive impairment (Bassuk and Berkman, 1999). A recent cross-sectional study demonstrated that intergenerational support improved the cognitive functioning of the elderly in rural China (Wang, Li and Zhang, 2005). However, within the context of high levels of migration of working-age adults from rural to urban areas of China, multivariate analysis of longitudinal data on changes in the cognitive functioning of the elderly is required. This article presents an analysis of selected data from a 3-wave longitudinal survey conducted from 2001 to 2006 among 997 parents aged 60 and older living in rural Anhui Province, China, showing that intergenerational transfers, especially financial transfers, have a significant influence on the changes in the cognitive functioning of the elderly. As such, the findings bring a new dimension to the current scientific discussion on the impacts of changing traditional patterns of intergenerational support for elderly persons in China.

Studies in China have generally found that social support, particularly emotional support from relatives, is beneficial for the psychological well-being of older persons (He, 2002; Chen and Silverstein, 2000). The evidence is more equivocal regarding the benefits of instrumental support. While studies of older adults in Beijing (Chen and Silverstein, 2000) and Shanghai (Cui and Li, 1997) suggest that instrumental support may compromise well-being by inducing dependency, evidence from Hong Kong, China shows that both material and instrumental support were more influential than emotional support in preventing depression in older adults (Chi and Chou, 2001). A study in Wuhan city found that receiving financial support from informal sources was detrimental to the psychological well-being of older adults, presumably because the reallocation of scarce resources caused strain in these supportive relationships (Krause, Liang and Gu, 1998). The effects of intergenerational transfers – financial, emotional and instrumental – may be more pronounced for older people in rural areas, where filial norms are generally stronger and where the family support systems are more institutionalized (Davis-Friedmann, 1991).

Some research has found that the migration of working-age adults from rural to urban areas of China has resulted in greater geographic separation between the generations, reducing opportunities for older people to live with their children in a stable home environment, and has hindered intergenerational transfers (Rossi and Rossi, 1990). However, other research has found that there have been some differences in the impacts of the three types of intergenerational support. Close residential proximity of people of different generations has not only contributed to intergenerational support, especially instrumental support, but has also contributed to emotional closeness between the generations (Crimmins and

Ingegneri, 1990). Litwark and Kulis (1987) argued that the evidence that geographic separation affected financial and emotional support was weak.

The economic improvement that results from the migration of adult children can increase financial support to the elderly from adult children, contributing to significantly improving the health and well-being of the elderly in China (Rogers, 1996). Baseline data of this longitudinal study suggest that financial transfers could improve the cognitive functioning of the elderly in rural China (Wang, Li and Zhang, 2005). However, further analysis of the longitudinal data is required to examine whether increasing financial support affects changes in the cognitive functioning of the elderly.

Previous studies have shown that the migration of adult children places greater burdens upon the elderly. Elderly persons whose adult children have migrated may have to care for their grandchildren, and there are also clear negative psychological and physical consequences for many elderly women as a result of the migration of their children (Alspaugh, Zarit and Greene, 1999; Son, Zarit and Rovine, 2003). Increased emotional burdens may result in cognitive impairment of the elderly. Another study indicated that elderly persons in good health who have received more instrumental support face more risks than those who received less instrumental support (Seeman, Bruce and McAvay, 1996), with the result being harmful changes in the cognitive functioning of the elderly.

The "use it or lose it" hypothesis suggests that mental stimulation can protect cognitive faculties (Holtzman and others, 2004). In rural China, the limited social activities of the elderly result in them relying on their children for mental stimulation (Zhang, 2001), and emotional transfer between generations contributes to maintaining the cognitive functioning of the elderly (Wang, Li and Zhang, 2005). Reduced opportunities for interaction between generations, resulting from the migration of adult children, can therefore lead to a reduction in the level of cognitive functioning of the elderly. However, improvements in transportation and communication networks break spatial isolation, and may assist in emotional transfers and hence contribute to the cognitive functioning of the elderly.

There are many sociodemographic factors that have also been shown to be associated with the cognitive functioning of the elderly. These variables include age, household income, health status (including the initial status of cognitive functioning), sex, education, occupation and marital status, all of which have different influences on cognitive functioning (Hultsch and others, 1998; Jacqmin-Gadda and others, 1997; Lee and others, 2003; MacKnight and others, 2002; Ross and Mirowsky, 1999; Xu, 2001). Any study that attempts to assess the extent to which intergenerational transfers contribute to changes in the cognitive

functioning of the elderly needs to control for these confounding factors. The three-wave survey from which the data discussed in this article derive complies with this condition, having not only measured changes in the levels of intergenerational support, but also having collected data on both time-varying and time-invariant sociodemographic and economic characteristics.

Research framework and methods

Data for this study were derived from the three-wave survey "Well-being of the elderly in Anhui Province, China", conducted in Chaohu city of Anhui Province in April 2001, in November 2003 and in December 2006 by the Institute for Population and Development Studies of Xi'an Jiaotong University.

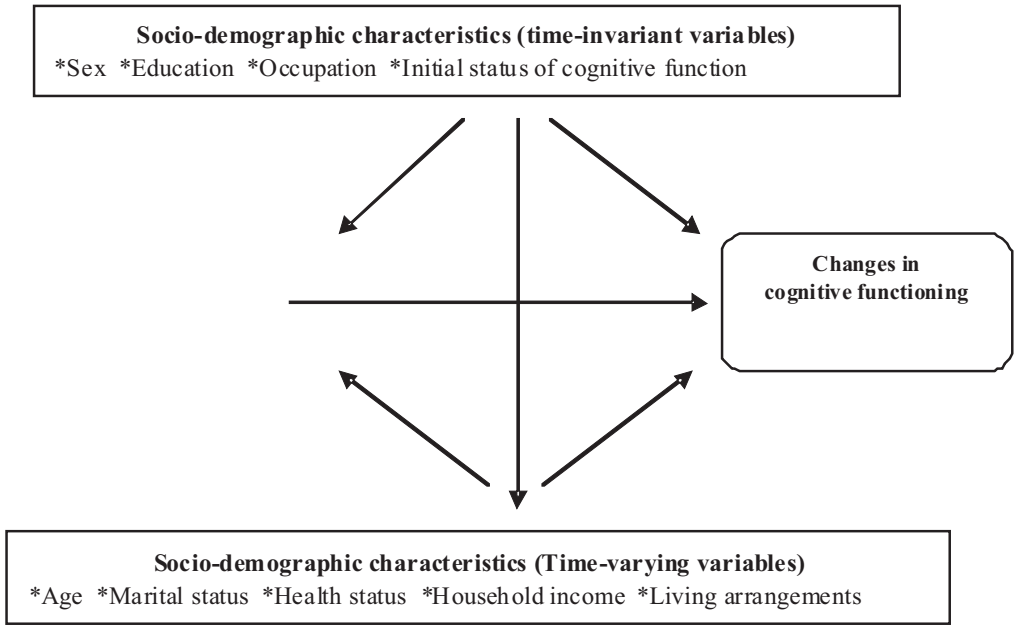
The survey included assessments of family relations, physical health status and psychological well-being. The respondents were selected from all residents aged 60 and older in the rural areas of Chaohu. Stratified multistage sampling was adopted to randomly select 1,800 respondents; of that number, 1,715 completed the survey, yielding a response rate of 95.3 per cent for 2001. Follow-up surveys were conducted in 2003 and 2006, with 1,391 and 1,067 respondents of the original survey completing the surveys, producing response rates of 81.11 and 76.71 per cent respectively. Of those respondents, who could not be located, 69 and 57 had moved out of their village and 240 and 236 died, respectively. There were 15 and 31 former respondents who could be located but refused to participate, terminated their interviews, or were too ill to be interviewed. After deleting those with missing values in the relevant variables, we restricted our analysis to the 997 elderly parents who had at least one living child.

The data produced were used in the study presented here to examine the direct and indirect impacts of intergenerational transfers on changes in cognitive functioning of the elderly in rural China. The conceptual framework of the study is shown in figure 1. Based on the principle of reciprocity, the study focuses on receiving and providing three types of intergenerational support: financial support, instrumental support and emotional support. The study addresses the following research hypotheses:

- H1: Increased financial support from, or provided to, adult children that contributes to changes in the cognitive functioning of the elderly in rural China;
- H2: Increased instrumental support from, or provided to, adult children that contributes to changes in the cognitive functioning of the elderly in rural China;

- H3: Increased emotional support between the elderly and adult children that contributes to changes in the cognitive functioning of the elderly in rural China.

Figure 1. Research framework



The use of hierarchical linear models represents an alternative to ordinary least squares for estimating individual growth or change curves (Raudenbush and Bryk, 1992). A growth curve model is used to estimate the relationship between intergenerational support and the changes in cognitive functioning in this study. A growth curve model characterizes the study of growth as a two-stage process, involving the specification of predictors of change measured at the individual and group levels. The level 1 model for intra-individual change describes the underlying change for each person as a function of time and a set of model growth parameters that define the change function. The level 2 model for interindividual differences in change describes how differences in growth across persons are related systematically to differences in various predictors of growth. In general, growth models of intra-individual differences utilize the approach shown below. The level 1 model is:

$$Y_{ji} = \pi_{0i} + \pi_{1i}(Time_{ji}) + r_{ji}$$

Where an observation at time *j* for an individual is a function of systematic growth trajectory plus a random error representing deviation from the trajectory at a particular observation point. $Time_{ji}$ represents age from the first to the last test; π_{0i} the intercept, is the older person's status at the first test; π_{1i}, the slope, is the linear growth rating period *e* for person *i* over the testing period. The level 2 model explains interindividual differences, or variations among persons, in initial status and rates of change.

To best describe variability in individual change, change in the dependent variable is described first. However, in attempting to account for systematic variance in estimated parameters, it must be determined that significant variance does exist in those parameters (Karney and Bradbury, 1995). If any parameter does not vary across individuals, then that parameter need not be examined further in the second stage of the analysis, in which conditional models are estimated. Individual growth-curve models analyse changes in cognitive functioning by focusing on interindividual differences in intra-individual change. The structure of the mean growth trajectories is described in the fixed effects models. The extent to which individual variation around mean growth occurs is estimated by individual variability in initial status and rate of cognitive function decline. In such analyses, time is considered to be nested within the individual. The hierarchical linear model (HLM), which uses HLM version 6.02 software, is employed here for growth-curve modelling.

The dependent variable is the score of time-varying cognitive functioning that was measured in each of the three surveys. Each incident of cognitive functioning was measured with three performance tests designed to assess short-term and working memory. Short-term memory was assessed by the number of words out of five that the respondents correctly recollected 30 seconds after they had been read. Working memory was assessed by serial subtraction, that is, the number of times, starting at 20, that the respondents could successively subtract 3. Orientation to time and place was assessed by correct answers to five questions testing the respondents' knowledge of their current condition. After reverse coding the items, we used principal components analysis to create a standardized factor score that would denote the degree of cognitive impairment. The reliability coefficient for these items was 0.76. In summing the score of the six items, the total score is considered as the status of cognitive functioning. The value of the cognitive functioning variable is the sum of the scores of all 6 items in this measure, ranging from 7 to 19. A higher value of this measure reflects greater cognitive functioning.

Predictor variables were grouped into two categories: intergenerational support (providing and receiving) and sociodemographic characteristics. Intergenerational support was measured by financial support, instrumental support and emotional

support. Assessment of financial support from and to children was based on measures of the total amount that the parent received from each child during the previous 12 months. Response options were the following categories based on Chinese renminbi (RMB) currency (RMB 100 = US$7.50): 0 = none; 1 = less than RMB 50; 2 = RMB 50-99; 3 = RMB 100-199; 4 = RMB 200-499; 5 = RMB 500-999; 6 = RMB 1,000-2,999; 7 = RMB 3,000-4,999; 8 = RMB 5,000-9,999 and 9 = more than RMB 10,000. The logarithm of the sum of all financial transfers received from adult children in each survey is considered as the financial support received by the elderly. The financial support provided was measured by answers to the question: "Did you (or your spouse living with you now) send your child money, food or gifts?" The measurement of the financial support provided was the same as for the financial support received. We took the total value across all children to represent the financial assistance received and provided.

The instrumental support received from children (including from the children's spouse and grandchildren) by elderly parents during the 12 months prior to each survey was measured in two areas: (a) household tasks, such as cleaning the house and washing clothes, and (b) personal care tasks, such as assistance in bathing and dressing, which are classified into four scales: (1) every day = 4; (2) at least once a week = 3; (3) several times a month = 2 and (4) seldom = 1. In summing the frequency of the instrumental support received from each child, the total score is considered as the instrumental support received by the elderly people in each survey.

Emotional support was measured by three questions assessing the quality of parent-child relationships. Three questions were adapted from the intergenerational solidarity inventory (Mangen, Bengtson and Landry, 1988), which assesses emotional cohesion between generations: (a) "Taking everything into consideration, how close do you feel to (this child)?" (b) To what degree do you feel that (this child) would be willing to listen when you need to talk about your worries and problems? and (c) "Overall, how well do you and (this child) get along together?" The items were coded as follows: 0 = not at all close/not at all/not at all well; 1 = somewhat close/somewhat/somewhat well; 2 = very close/very much/very well. An additive scale was computed for each child, ranging from 0-6. We took the highest total score across all children for each parent to indicate this construct. The reliability coefficient for these items was 0.82.

Sociodemographic control variables included age, income, health status, sex, education, occupation, marital status and living arrangements. Age, income and health were time-varying variables measured at each of the three waves of the survey. The variable "age" was set as "age 60", which means that it is the person's age minus 60. Because instrumental activities of daily living rely greatly on cognitive

functioning, in this study health status was measured as the extent of functional problems, and the reliability of this scale was 0.88. The remaining variables were coded as dummy variables: sex (1 = female); martial status (1 = currently not married); education (1 = some formal education); past or current occupation (1 = agricultural work) and living arrangements (1 = living alone).

Results and discussion

The descriptive profile shown in table 1 provides the mean values of the dependent variable at the baseline and follow-up surveys by the sex of the respondent. The trend of cognitive functioning declined and the difference, especially in the standard deviation, between the individuals became larger over time. Cognitive functioning for elderly males was always higher than that for elderly females in the three-wave survey. As expected, women on average were older, less educated and more likely to not be currently married compared with men. Over the three rounds of the survey, the elderly, especially older women, were more likely to become single, live alone, have more functional health difficulties and suffer from depression, and receive more household income. In the three-wave survey, the elderly received more money from their adult children over time. Older women received more instrumental support from their children over the three waves, while older men received less. The elderly provided a declining amount of instrumental support to their children over the three waves.

Table 1. Summary statistics of study variables (N = 997)

Variable	Males		Females	
	Mean	SD	Mean	SD
Dependent (cognitive functioning)				
Wave 1 (2001)	15.74	2.32	13.85	2.47
Wave 2 (2003)	15.85	2.33	13.87	2.55
Wave 3 (2006)	15.12	2.60	13.14	2.84
Indeoendent variables				
Sociodemographic				
Age				
Wave 1	68.32	5.92	70.94	6.90
Wave 2	70.25	5.90	72.90	6.82
Wave 3	73.33	5.91	75.89	6.92
Education (wave 1) (1 = educated)	0.43	0.50	0.07	0.25
Occupation (wave 1) (1 = agricultural work)	1.00	0.00	1.00	0.00
Marital status (1 = not currently married)				
Wave 1	0.22	0.42	0.49	0.50
Wave 2	0.24	0.43	0.54	0.50
Wave 3	0.31	0.46	0.61	0.49

.../

Table 1 (*Continued*)

Variable	Males		Females	
	Mean	SD	Mean	SD
Functional health difficulties				
Wave 1	17.15	4.02	21.14	6.51
Wave 2	17.39	4.70	21.25	6.92
Wave 3	19.77	6.45	23.71	8.03
Depression (CES-D)				
Wave 1	14.39	3.74	15.68	3.77
Wave 2	13.96	3.68	15.67	3.9
Wave 3	14.96	4.05	16.53	3.80
Household income (in RMB)				
Wave 1	3 128.35	12 040.66	1 891.77	4 827.88
Wave 2	2 842.31	2 690.59	2 090.09	2 404.28
Wave 3	3 690.00	3 299.22	2 922.69	3 150.29
Living arrangements (1 = live alone)				
Wave 1	0.13	0.33	0.19	0.39
Wave 2	0.15	0.36	0.22	0.41
Wave 3	0.18	0.38	0.24	0.43
Intergenerational support				
Received money from children				
Wave 1	1 135.69	1 665.61	968.56	1 121.33
Wave 2	1 421.18	1 733.98	1 187.01	1 150.45
Wave 3	1 730.88	1 931.88	1 715.82	1 816.58
Provided children with money				
Wave 1	319.58	1 073.27	147.53	527.70
Wave 2	257.67	775.78	150.51	480.55
Wave 3	475.25	2 674.23	272.99	1 312.08
Received instrumental support from chilfren				
Wave 1	2.67	5.87	3.80	5.34
Wave 2	1.99	5.49	4.06	6.28
Wave 3	1.80	3.66	4.36	6.70
Provided children with instrumental support				
Wave 1	1.25	2.72	2.10	3.20
Wave 2	1.09	2.89	1.89	3.55
Wave 3	0.50	1.62	1.02	2.59
Emotional support transfer				
Wave 1	7.17	1.28	7.24	1.22
Wave 2	7.41	1.29	7.45	1.21
Wave 3	7.22	1.35	7.01	1.25
Number of cases		457		540

Source: Survey of "Well-being of the elderly in Anhui Province" conducted in Chaohu City, 2001, 2003 and 2006.

The results of the unconditional mean model and unconditional linear growth model of cognitive functioning over three time intervals are shown in model 1 and model 2 of table 2. The results indicate that there are significant developmental differences in the cognitive functioning of the elderly. Therefore, conditional models are tested and the results shown in table 3.

Table 2. Linear growth models and effect of cognitive function of the baseline (N = 997)

	Model 1		Model 2	
	Coefficient (SE)	T value	Coefficient (SE)	T value
Fixed effects				
INTRCPT:				
Initial status of cognitive function	14.474 (0.071)	203.960***	15.914 (0.119)	133.833***
Slope: Age			-0.119 (0.009)	-13.384***
Random effects	Variance (SD)	X^2	Variance (SD)	X^2
INTRCPT	3.85648 (1.963)	4 298.79***	2.79715 (1.672)	1 298.789***
Slope			0.00409 (0.064)	1 115.808**
Level-1, R	3.493 (1.869)		3.36361 (1.834)	

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

In table 3, the fixed effects model of model 1 examined the effects of intergenerational support on the changes in the elderly's cognitive functioning. Its results indicate that there was a significant decline in cognitive functioning associated with ageing (-0.105). There were significantly positive influences on the changes in cognitive functioning when the elderly received money from their adult children and provided their children with money (0.047, 0.074), while there was a significantly negative influence on the changes in cognitive functioning when the elderly received instrumental support from their children (-0.026). Emotional support between the elderly and their children improved cognitive functioning (0.066) over time. The random effects model of model 1 indicated that there were still significant differences in the individual intercept (2.59107) and slope (0.00357) after adding the types of intergenerational support. These mean that there were significant declines in cognitive functioning over the six-year period and that the rate of decline varied among individuals. Levels of intergenerational support explained almost 7.4 per cent of the initial status of cognitive functioning (2.79715-2.59107)/2.79715 = 7.37 per cent, and it explained

12.7 per cent of the rate of decline of cognitive functioning, i.e. (0.00409-0.00357)/0.00409 = 12.71 per cent.

Table 3. Hierarchical linear model predicting cognitive functioning during six years of follow-up (N = 997)

	Model 1		Model 2	
	Coefficient (SE)	**T value**	**Coefficient (SE)**	**T value**
Fixed effects				
Initial status of cognitive functioning	14.864 (0.316)	46.996***	16.949 (0.395)	42.881***
Age	-0.105 (0.009)	-11.386	-0.051 (0.009)	-5.550***
Sex (wave 1) (1 = male)			0.875 (0.127)	6.913***
Education (wave 1) (1 = educated)			1.194 (0.133)	9.012***
Occupation (wave 1) (1 = agricultural work)			-0.368 (0.236)	-1.561
Marital status (1 = not married)			-0.088 (0.128)	-0.689
Household income			-0.000 (0.000)	-0.158
Functional health difficulties (IADL)			-0.187 (0.019)	-9.677***
Depression (CES-D)			-0.060 (0.012)	-4.909***
Living arrangement (1 = living alone)			0.193 (0.131)	1.474
Intergenerational support				
Received money from children	0.047 (0.027)	1.737^{+}	0.053 (0.025)	2.08*
Provided children with money	0.074 (0.016)	4.552***	0.050 (0.015)	3.239**
Received instrumental support from children	-0.026 (0.008)	-3.193**	0.001 (0.008)	0.121
Provided children with instrumental support	0.019 (0.015)	1.270	0.016 (0.014)	1.077
Emotional support transfer	0.066 (0.037)	1.794^{+}	-0.008 (0.036)	-0.215
Random effects	Variance (SD)	X^2	Variance (SD)	X^2
INTRCPT	2.59107 (1.610)	1 289.440***	1.66692 (1.291)	1 216.497***
Slope	0.00357 (0.060)	1 118.02**	0.00254 (0.050)	1 105.412**
Level-1	3.385 (1.840)		3.325 (1.823)	

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; $^{+}$ $p < 0.1$

In model 2 of table 3, when we added sociodemographic variables, the results changed. For the slope of level 1, the fixed effects model of model 2 indicated that there were significantly positive influences on the changes in cognitive functioning when the elderly received money from their children and provided their children with money (0.053, 0.050), but the significant influences on the changes in cognitive functioning that were derived from other forms of intergenerational support were no longer significant. The mean of cognitive initial status for the men or the educated was significantly better than their respective counterparts. Cognitive functioning declined with the ageing of the elderly (-0.051), when they had more functional health difficulties (-0.187) and when they experienced worsening well-being Center for Epidemiologic Studies Depression Scale (CES-D)(-0.060). There was no significant influence on the changes in cognitive functioning associated with whether or not the elderly engaged in agricultural work, were currently married, had a high or low household income, or were living alone. The random effects model indicated that there were still significant differences in the individual intercept (1.66692) and slope (0.00254) after controlling for sociodemographic characteristics, that is to say, there was a significant decline in cognitive functioning over the six-year period and the rate of decline varied among individuals. Intergenerational support explained almost 35.7 per cent of the initial status of cognitive functioning, i.e. (2.59107-1.66692)/ 2.59107 = 35.67 per cent, and it explained 28.9 per cent of the rate of decline in the cognitive functioning level (0.00357-0.00254)/0.00357 = 28.85 per cent.

Conclusion

Rapid economic growth, accelerating migration patterns and changing family preferences in China have jeopardized the ability of older parents to rely on the adult children for needed support and this situation is compromising the well-being of the elderly. In this study, data derived from a three-wave longitudinal survey from 2001 to 2006 of 997 parents aged 60 and older living in rural Anhui Province are analysed. The hierarchical linear model was used to estimate the effects of the time-varying variables, which are intergenerational transfers of financial, instrumental and emotional support, on the changes in cognitive functioning of the elderly living in a highly mobile area of rural China.

It was found that there is significant variability in the initial status or rate of change in cognitive functioning of the elderly at any interval. On average, the changes in cognitive functioning decreased over time. In addition, there is a positive relationship between increasing levels of receiving and providing money and changes in cognitive function, even after controlling for the

sociodemographic characteristics of the elderly. The results support the first hypothesis of the study and partly support the second hypothesis. The negative relationship between increasing levels of receiving instrumental support from adult children and changes in cognitive functioning of the elderly in model 1 supported the second hypothesis. However, after controlling for the needs of the elderly, a positive relationship was found between increasing levels of receiving instrumental support from adult children and the changes in the cognitive functioning of the elderly. The third hypothesis was also partly supported. The significant effect of emotional support shown in model 1 of table 3 supports the third hypothesis, but the statistical significance of the relationship disappeared in model 2 of table 3. This result may be due to increasing levels of functional health difficulties, a situation which leads to the need for the elderly to receive more instrumental support and results in a worsening of their psychological well-being. Such a situation could increase conflicts with caregivers owing to the greater burden being placed on them and further the decline in the cognitive functioning of the elderly.

The control variables used in the analysis, such as ageing, female sex, being uneducated and having functional health difficulties and poorer well-being, were all significant risk factors for changes in the cognitive functioning of the elderly. Agricultural work did not have a significant influence, perhaps because most respondents were engaged in this activity. The reason why marital status had no significant effects may be that it affects cognitive functioning of the rural elderly indirectly through intergenerational support, with adult children being more likely to provide support for elderly parents without a spouse (Xu, 1996). There is a direct relationship between living alone and changes in cognitive functioning, with the results showing that living alone contributes to maintaining the activities of daily living of the elderly and lessens the decline in the psychological ability associated with ageing (Buer and Linton, 2002).

The present study has several strengths, including the use of longitudinal data and their enabling study of time-bound trends. The analytic technique utilized here provides further understanding of how cognitive functioning changes over time. The data, from a rural area sample of the elderly in China, also enable a determination of the effect of intergenerational support on the changes in cognitive functioning. However, there is also a need to consider several limitations of the data and analyses. Only three waves of data are currently available. Additional waves of data would lead to higher precision for estimating the individual growth trajectory and greater reliability in the measurement of change. Because the increasing rate of cognitive impairment among the elderly in

rural China poses significant problems, not only for the individuals and families concerned but for Chinese society as a whole, there is a need to urgently develop policies and programmes that will reduce the impact of the negative changes and maximize that of the positive changes.

References

Alspaugh, M.E.L., S.H. Zarit and R. Greene (1999). "Longitudinal patterns of risk for depression in dementia caregivers: Objective and subjective primary stress as predictors", *Psychology of Aging,* vol.14, No.1, pp. 34-43.

Bassuk, S.S, T.A., Glass and L.F. Berkman (1999). "Social disengagement and incident cognitive decline in community-dwelling elderly persons", *Annals of Internal Medicine*, vol. 131, pp. 165-173.

Bradbury, T.N. and others (1995). "Assessing longitudinal change in marriage: An introduction to the analysis of growth curves", *Journal of Marriage and the Family,* vol. 57, No. 4, pp. 1091-1108.

Bryk, A.S. and S.W. Raudenbush (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*, Thousand Oaks*,* California, United States, Sage Publications, Inc.

Buer, N. and S.L. Linton (2002). "Fear avoidance beliefs and catastrophizing: Occurrence and risk factor in back pain and ADL in the general population", *Pain*, vol. 99, No. 3, pp. 485-491.

Chen, X. and M. Silverstein (2000). "Intergenerational social support and the psychological well-being of older parents in China", *Research on Aging,* vol. 22, No. 1, pp. 43-65.

Chi, I. and K.L. Chou (2001). "Social support and depression among elderly Chinese people in Hong Kong, *International Journal of Aging and Human Development*, vol. 52, No. 3, pp. 231-252.

Crimmins, E.M. and D.G. Ingegner (1990), "Interaction and living arrangements of older parents and their children: Past trends, present determinants, future implications", *Research on Aging*, vol.12, No. 1, pp. 3-35.

Cui, L. and H. Li (1997). "Urban elders' social support network and life satisfaction", *Psychological Science,* vol. 20, pp. 123-126.

Davis-Friedmann, D. (1991). *Long Lives: Chinese Elderly and the Communist Revolution,* 2nd ed*.,* Cambridge, Massachusetts, United States, Harvard University Press.

He, Z. (2002). "Socioeconomic status and social support network of the rural elderly and their physical and mental health", *Social Science in China*, vol. 3, pp.135-148.

Holtzman. R.E., G.W. Rebok, J.S. Saczynski, A.C. Kouzis, K. Wilcox Doyle and W.W. Eaton (2004). "Social network characteristics and cognition in middle-aged and older adults", *The Journals of Gerontology Series B: Psychological Sciences*, vol. 59, pp. 278-284.

Hultsch, David F., C. Hertzog, R.A. Dixon and B.J. Small (1998). *Memory Change in the Aged,* Cambridge, Massachusetts, United States, Cambridge University Press.

Jacqmin-Gadda, H., C. Fabrigoule, D. Commenges and J.F. Dartigues (1997). "A 5-year longitudinal study of the Mini-Mental State Examination in normal aging", *American Journal of Epidemiology,* vol. 145, No. 6, pp. 498-506.

Krause, N., J. Liang and S. Gu (1998). "Financial strain, received support, anticipated support, and depressive symptoms in the People's Republic of China", *Psychology and Aging,* vol. 13, pp. 58-68.

Lee, S., I. Kawachi, L.F. Berkman and F. Grodstein (2003). "Education, other socioeconomic indicators, and cognitive function", *American Journal of Epidemiology,* vol. 157, No. 8, pp. 712-720.

Litwark, E. and S. Kulis (1987). "Technology, proximity and measurement of kin support", *Journal of Marriage and the Family,* vol. 49, pp. 649-661.

MacKnight, C., K. Rockwood, E. Awalt and I. McDowell (2002). "Diabetes mellitus and the risk of dementia, Alzheimer's disease and vascular cognitive impairment in the Canadian Study of Health and Aging", *Dementia and Geriatric Cognitive Disorder*, vol. 14, pp. 77-83.

Rogers, R.G. (1996). "The effects of family composition, health and social support linkages on morality", *Journal of Health and Social Behavior*, vol. 37, No. 4, pp. 326-338.

Ross, C.E. and J. Mirowsky (1999). "Refining the association between education and health: The effects of quantity, credential, and selectivity", *Demography,* vol. 36, No. 4, pp. 445-460.

Rossi, A.S. and P.H. Rossi (1990). *Of Human Bonding: Parent-Child Relations across the Life Course,* New York, Aldine de Gruyter.

Seeman, T. E., M.L. Bruce and G.J. McAvay (1996). "Baseline social network characteristics and onset of ADL disability: MacArthur Studies of Successful Aging", The *Journals of Gerontology Series B: Social Sciences,* vol. 51, pp. S191-S200.

Son, J., S. Zarit and M. Rovine (2003). "Caregiving stress process: Social support as a mediator", *Gerontologist,* vol. 43, No. 1, pp. 43.

Wang, P., S. Li and W. Zhang (2005). "The effects of intergenerational supports on the cognitive function of the rural elderly in China" (in Chinese), *Psychological Science*. vol. 28, No. 6, pp.1500-1503.

Xu, Q. (1996)."Comparative study on the intergenerational support of sons and daughters (in Chinese), *Journal of Population Research*, vol. 5, pp. 23-31.

_____ (2001). "Analysing the mental status of the oldest old" (in Chinese), *Population Journal*, vol. 19, No. 5, pp. 45-52.

Zhang Y. (2001). "Comparative study of older people's social support networks in rural and urban areas of China – A case study in Xiamen" (in Chinese), *Sociology Research,* vol. 4, pp. 11-21.

Evaluation of the Thai Tobacco Control Policy

Convincing smokers of the need to quit is best done by emphasizing the negative impacts of smoking on health, not only their own but also the health of others, especially family members.

By Buppha Sirirassamee, Philip Guest, Warangkana Polprasert, Tawima Sirirassamee, Orapin Pitakmahaket and Pariya Gainroj*

Tobacco use has been identified by the World Health Organization (WHO) as the most preventable cause of death and disability in the world. Globally, there are 1.1 billion people who smoke, over 80 per cent of whom live in low- and middle-income countries. Of all the people alive today, 500 million will die of tobacco-related causes. Although a staggering 100 million tobacco-related deaths occurred in the twentieth century, it is estimated that, by the end of the twenty-first century, 10 times that number, or 1 billion people, will have died because of tobacco use, and this burden will be borne most heavily by developing countries.

* Buppha Sirirassamee, Philip Guest, Orapin Pitakmahaket and Pariya Gainroj, Institute for Population and Social Research, Mahidol University; Warangkana Polprasert, Sukhothai Thammathirat Open University; and Tawima Sirirassamee, Faculty of Medicine, Srinakharinwirot University, Thailand. E-mail: prbsr@mahidol.ac.th.

Currently, 5 million people die of tobacco use each year worldwide, but by 2030, the number of such deaths will rise to 10 million per year and 70 per cent of them will be in developing countries. In addition to the personal costs of this preventable tragedy will be the economic and social costs; these will pose major challenges to the aspirations of greater economic growth and stability in the developing world (WHO, 1997).

In recognition of the threat that tobacco use signifies, all 192 WHO Member States, representing 95 per cent of the world's population, recently adopted the WHO Framework Convention on Tobacco Control, the first treaty devoted entirely to health, and over 120 countries have already ratified it. The convention specifies national-level tobacco-control policies that the signatories are or will be obligated to implement in the coming years. These include banning of misleading brand descriptors, restrictions/prohibitions on advertising and the promotion of tobacco products, increases in taxation, measures to limit exposure to second-hand smoke and those designed to eliminate illicit trade in tobacco products.

The convention is a major milestone in global health, yet there are formidable challenges in approaching the implementation phase, especially that of facilitating the passage of the convention's provisions in a timely fashion. Countries throughout the world will face extraordinary resistance in this process from a variety of quarters, particularly the tobacco industry. Tobacco control policymakers will need to compile and put forward their best case for implementation of the convention's policies. It is in this context that evidence supporting the efficacy of the convention's policies will be in great demand.

As tobacco-control policies are formulated and implemented, it is important for such policies to be accompanied by rigorous evaluation. Good evaluation not only answers the question of whether a policy works as intended, but it also should demonstrate the mechanisms by which it is effective and the reasons for any lack of success. Generally speaking, a policy may not work as intended for three broad classes of reasons, as described below.

First, the policy is wrong. The behaviours it seeks to change do not actually affect tobacco use as hypothesized, or the policy does not have the capacity to change the target behaviours. This may be an inherent limitation of the policy or a failure in some socio-cultural contexts. For example, a policy linking tobacco use to religious concerns may be very effective in a country or region which follows a particular religion, but may have no effect in countries with different religions or in secular societies.

Second, a policy may not work because of failures in implementation. Two types of failure are particularly important in this context: failure to provide the

necessary infrastructure to support implementation and failure to adapt the policy or programme to the socio-cultural context.

In tobacco control there is a third possible class of reasons: countervailing actions by the tobacco industry. The tobacco industry has the capacity to modify its products and their distribution as well as the marketing of those products.

In discerning the various patterns of implementation, longitudinal data are particularly important. The collection of data from the same population, indeed the same persons, before and after the introduction of the policy provides the data that will enable assessment of the impact of tobacco-control policies. In particular, it will enable assessment of the extent to which individuals have been exposed to the tobacco-control policies and how such exposure has had an impact upon their behaviour, as demonstrated in this article which presents the initial findings from an ongoing study in Thailand.

Thailand has joined the longitudinal International Tobacco Control Policy Evaluation Project (ITC Project) in order to examine the impact of its tobacco-control policies. As described by Thompson and others (2006), the ITC Project has two key design features. The first of these involves is the use of natural experiments. In essence, this means that tobacco controls are occurring in many countries and study designs can take advantage of these controls by designing data-collection schemes that provide evidence of the impact of the controls. In turn, this relates to the second design feature of the study, the collection of longitudinal data with a panel design. In this design, the respondents are differentially exposed over time to policy changes, which provides researchers with the opportunity to differentiate changes over time within individuals from differences in the baseline levels, or cohort effects.

Study background and methods

The prevalence of smoking tobacco among Thais has declined since 1976. The surveys conducted by the National Statistical Office (NSO) from 1976 to 2006 found that in 1976 the prevalence of smoking among Thais aged 15 and older was 30.1 per cent, slowly declining to 26.4 per cent in 1986. The latest survey in 2006 found that 18.9 per cent of Thais aged 15 and older were smokers (table 1). Males are much more likely to smoke than females. The average age at which a person smokes his or her first cigarette is 18 years old, and the average number of cigarettes smoked per day is 10.4 (NSO, 2006).

Thailand is widely seen as a global leader in tobacco control and thus the tobacco-control policies and programmes that are implemented in Thailand are of

considerable interest and importance throughout the world. This is particularly true among developing countries. Thailand has recently implemented two world-leading tobacco control policies: the introduction of graphic warning labels on cigarette packaging in March 2005 and the ban on all marketing (signage, products) at point-of-sale locations in September 2005, both of which are anti-tobacco measures recommended by the convention. No studies have evaluated the impact of graphic warnings in low- or middle-income countries. Evidence concerning the Thai graphic warning labels would thus be crucial in informing labelling policies throughout the world. The evaluation of the Thai point-of-sale ban would be important because it would be the first evaluation of a national comprehensive point-of-sale ban.

Table 1. Number and percentage of smokers aged 15 and older, 1976-2006

Year	Number	Percentage
1976	8 629 510	30.1
1986*	10 377 000	26.4
1988	10 109 890	25.0
1991	11 402 100	26.3
1993*	10 406 200	22.8
1996	11 254 200	23.4
1999	10 230 600	22.4
2001	10 570 100	22.5
2004	9 627 600	19.5
2006	9 507 000	18.9

Source: National Statistical Office (2006).
* Aged 10 years and older.

To assess the extent of the introduced measures and their impact on people's behaviour, a study was initiated as part of the South East Asia Tobacco Control Policy Survey-ITC-SEA (Thailand). That study, which adopted a prospective cohort design, will last for five years (2005 to 2009). It compares the results of the first and second waves of a prospective study of adult smokers. The study has three major objectives, as follows:

(a) Explore smoking behaviour, the knowledge of, and attitudes towards smoking, and knowledge of tobacco policies;

(b) Analyse the individual and environmental determinants of smoking behaviour;

(c) Evaluate the effects of the tobacco-control policy on smoking behaviour.

The wave 1 survey (2005) involved 2,000 adult smokers (aged 18 years or older) who were sampled and interviewed in the period January-February 2005. Only those who smoked were included in the sample. Smokers were defined as those persons who had smoked at least 100 cigarettes in their lifetime and currently smoked at least 1 cigarette per week. Both females and males were eligible for sampling and had equal probabilities of selection, although only one male and one female respondent could be selected from a household. The sample was designed to be representative at the regional level and for rural and urban areas. The sample design was stratified multi-stage sampling.

After the wave 1 data collection was completed, the Government of Thailand implemented several new policies designed to reduce the incidence of smoking. These policies included the introduction of new warning displays on cigarette packets that were illustrated by graphic pictures of the damage to health caused by smoking, and the banning of displays of cigarettes at points of sale. A comparison of wave 1 and wave 2 data provides an opportunity to examine the impact of these new policies.

In the wave 2 survey (2006), the study attempted to interview all those respondents who were interviewed in wave 1. A total of 1,568 respondents were successfully reinterviewed in the period August-September 2006, resulting in a follow-up rate of 78.4 per cent. Only in Bangkok was the follow-up rate below 50 per cent (47.5 per cent). A replenishment sample of 512 adult smokers was also selected and interviewed. In this article, the results are provided only for the 1,568 respondents who were interviewed in both waves of the study. Where comparisons were made between the two waves, the analysis treats each wave as a separate sample.

Main findings

Of the 1,568 persons reinterviewed in wave 2, 1,358 (87.2 per cent) were smokers at wave 2 and the remaining 200 (12.8 per cent) had quit smoking in the interval between wave 1 and wave 2. Males comprised 90.9 per cent of the sample. The proportion of males was similar for the group that still smoked and the group that had quit smoking. The average age of smokers was 49 years and for those who had quit smoking the mean age was 53 years. Almost 75 per cent of the sample had a secondary school level of education and approximately 80 per cent of them were married.

Among the smokers, approximately half reported that their health was satisfactory. By comparison, almost 75 per cent of those who had quit smoking stated that their health was good or very good. Most smokers lived in households where at least one household member smoked, while most persons who had quit smoking lived in households where no one in the household smoked.

As can be seen from the data presented in table 2, the average number of cigarettes smoked daily declined significantly from 13.3 in wave 1 to 10.2 in wave 2. This decline was observed in both rural and urban areas. Among those persons who were smokers in both wave 1 and wave 2, the decline in the average number of cigarettes smoked daily was from 13.6 to 10.8. Much of the decline can be attributed to a high proportion of the smokers reducing their frequency of smoking from daily to less than daily. Apart from this change in frequency was a trend towards fewer cigarettes being smoked by those who smoked daily.

Table 2. Percentage distribution and mean of average number of cigarettes smoked per day, by place of residence and survey wave

Number of cigarettes smoked daily	Urban		Rural		Total	
	Wave 1 (406)	Wave 2 (355)	Wave1 (1,100)	Wave 2 (991)	Wave 1 (1,506)	Wave 2 (1,346)
0	0.5	8.5	0.0	13.5	0.1	12.2
1-4	8.6	14.9	11.4	14.1	10.6	14.3
5-8	16.3	19.2	17.4	18.1	17.1	18.4
9-12	33.3	25.4	27.1	24.8	28.8	25.0
13-16	10.8	9.6	13.6	11.5	12.9	11.0
17-20	24.6	16.3	22.1	12.7	22.8	13.7
21 or more	5.9	6.2	8.5	5.2	7.8	5.5
Total	100.0	100.0	100.0	100.0	100.0	100.0
Mean	13.2	10.9	13.3	9.9	13.3	10.2

Note: The 166 persons who did not smoke on a daily basis are coded as 0. The results of a paired t-test show that smokers decreased the daily number of cigarettes smoked from 13.6 to 10.8 (significant at the 0.0000 level).

The data for urban, rural and total are significant at the 0.001 level.

Also, there was a statistically significant shift between waves in the type of tobacco product smoked. In wave 1, more than a third (34.7 per cent) of the respondents smoked hand-rolled cigarettes only; this proportion had increased to half (49.9 per cent) in wave 2 (see table 3). This change is probably a result of the tax increases implemented by the Government, which increased the price of factory-produced cigarettes over that of hand-rolled ones.

The shift in smoking hand-rolled cigarettes was most apparent in rural areas, which may be because rural smokers are more sensitive to price increases than are urban smokers. The shift in type of cigarette smoked is not a function of the quitters being more likely to be smokers of factory-made cigarettes; the shift also occurred for those who were smoking in both waves (results not shown). The shift to hand-rolled cigarettes was noted mainly among those who previously had smoked both hand-rolled and factory-made cigarettes.

Table 3. Percentage distribution of main type of cigarettes smoked, by place of residence and survey wave

Type of cigarette	Urban		Rural		Total	
	Wave 1 (420)	Wave 2 (355)	Wave 1 (1,138)	Wave 2 (993)	Wave 1 (1,558)	Wave 2 (1,351)
Factory-made only	67.4	64.0	33.0	29.1	42.2	38.3
Hand-rolled only	19.3	27.9	40.4	57.8	34.7	49.9
Both types	13.3	8.1	26.6	13.1	23.0	11.8
Total	100.0	100.0	100.0	100.0	100.0	100.0

Note: The data for urban, rural and total are significant at the 0.001 level, and the numbers of smokers are contained within parentheses.

From wave 1 to wave 2 there were increases in the proportions of respondents who supported bans on smoking in a variety of public places, such as hospitals, air-conditioned and non-air-conditioned restaurants, workplaces, places of worship and public transport (see table 4). This phenomenon indicates that the general public increasingly accepted the restrictions on smoking.

There is an increasing belief that smoking has a strong impact on health, with almost 75 per cent of the respondents in wave 2 agreeing that smoking had a large impact on their health. This increase is largely confined to rural areas (see table 5). In wave 1 only 8 per cent reported that smoking had no effect on their health at all and this was reduced to 6 per cent in wave 2.

From table 6 it can be seen that knowledge of the adverse impacts of tobacco use on health increased between wave 1 and wave 2, and that knowledge is possibly related to the increasing proportions who believe that smoking has had a significant impact on health. Most of the increase was confined to conditions where knowledge was only moderate during wave 1: stroke and male impotence; statistically significant increases were also observed for knowledge about tobacco's effects on premature ageing. The increases are likely to be related to the pictorial warnings that have been introduced on cigarette packets.

Table 4. Percentage distribution of opinions of where smoking should be allowed in specified public places, by place of residence and survey wave

Areas where smoking should be allowed	Urban		Rural		Total	
	Wave 1 (420)	Wave 2 (414)	Wave 1 (1,137)	Wave 2 (1,137)	Wave 1 (1,557)	Wave 2 (1,551)
Hospital						
All indoor areas	0.5	0.0	0.2	0.5	0.3	0.4
Some indoor areas	22.4	10.9	28.0	15.8	26.5	14.5
Nowhere at all	77.1	89.1	71.8	83.7	73.2	85.2
Workplace						
All indoor areas	3.6	0.5	3.6	0.4	3.6	0.4
Some indoor areas	43.3	29.1	41.0	26.1	41.7	26.9
Nowhere at all	53.1	70.5	55.4	73.6	54.7	72.7
Air-conditioned restaurant						
All indoor areas	0.5	0.0	0.1	0.6	0.2	0.5
Some indoor areas	28.8	10.2	24.7	10.0	25.8	10.1
Nowhere at all	70.7	89.8	75.2	89.4	74.0	89.5
Non-air-conditioned restaurant						
All indoor areas	7.4	8.8	12.4	6.2	11.0	6.9
Some indoor areas	68.8	48.5	65.1	52.3	66.1	51.3
Nowhere at all	23.8	42.7	22.5	41.5	22.8	41.8
Public transport						
All indoor areas	0.5	0.2	0.0	0.8	0.1	0.6
Some indoor areas	0.7	2.2	3.0	1.9	2.4	2.0
Nowhere at all	98.8	97.6	97.0	97.3	97.5	97.4
Place of worship						
All indoor areas	2.1	1.0	3.3	0.9	3.0	0.9
Some indoor areas	43.8	4.6	47.7	4.7	46.7	4.7
Nowhere at all	54.0	94.4	48.9	94.4	50.3	94.4

Note: The data for all the locations are significant at the 0.001 level, except for the category public transport; the number of respondents is contained within parentheses.

Table 5. Percentage distribution of the self-reported effect of smoking on the health of the respondents, by place of residence and survey wave

Effect on health	Urban		Rural		Total	
	Wave 1 (417)	Wave 2 (404)	Wave 1 (1,128)	Wave 2 (1,123)	Wave 1 (1,545)	Wave 2 (1,527)
None at all	7.9	9.4	8.2	5.3	8.2	6.4
Some effect	26.6	21.8	28.3	18.9	27.8	19.6
Large effect	65.5	68.8	63.5	75.8	64.0	73.9
Total	100.0	100.0	100.0	100.0	100.0	100.0

Note: The data for rural and total are significant at the 0.001 level; the number of respondents is contained within parentheses.

Table 6. Percentage who agreed that smoking caused specified health outcomes, by place of residence and survey wave

Health outcome	Urban		Rural		Total	
	Wave 1 (420)	Wave 2 (413)	Wave 1 (1,138)	Wave 2 (1,140)	Wave 1 (1,558)	Wave 2 (1,553)
Stroke	40.2	73.1 ***	44.2	79.6 ***	43.1	79.6 ***
Male impotence	68.3	74.1 ***	69.9	76.9 ***	69.5	76.9 ***
Lung cancer in smokers	94.8	94.2	94.4	95.0	94.5	95.0
Stained teeth	93.1	95.7	95.1	95.3	94.5	95.3
Premature ageing	87.1	93.8 ***	88.0	91.6 ***	87.7	91.9 ***
Lung cancer in non-smokers	91.0	91.6	90.3	90.2	90.5	92.0

Note: The number of respondents is contained within parentheses.

*** Significant at the 0.001 level.

In comparing wave 1 and wave 2, there was a significant increase in the proportions who had noticed the warning labels on cigarette packets often or very often (see table 7). However, this increase was confined mainly to rural areas. The gap between rural and urban areas in the proportions who had noticed warning labels often, or very often, narrowed over the two waves. In wave 1, the gap was about 7 percentage points. By wave 2, the gap was less than 4 percentage points.

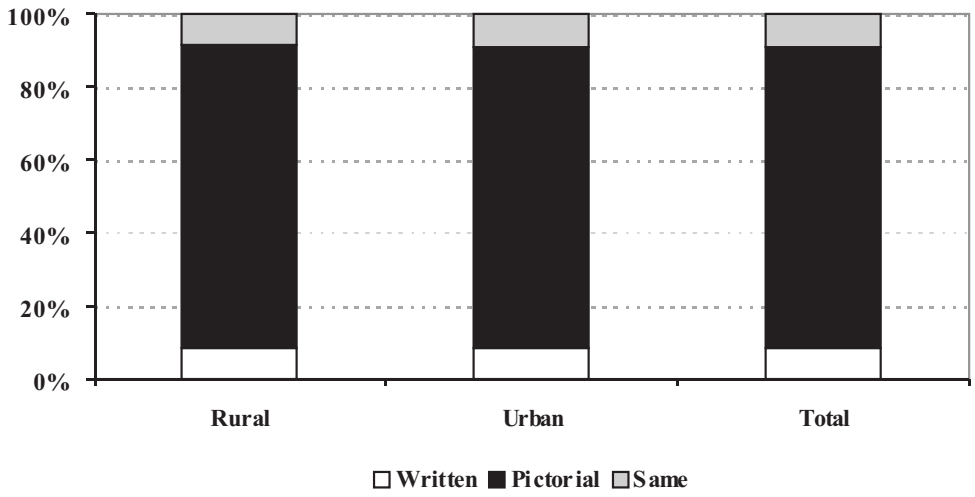
Table 7. Percentage distribution of the frequency of noticing warning labels on cigarette packets in the previous month, by place of residence and survey wave

Frequency of noticing warning labels	Urban		Rural ***		Total ***	
	Wave 1 (407)	Wave 2 (408)	Wave 1 (1,083)	Wave 2 (1,133)	Wave 1 (1,490)	Wave 2 (1,541)
Never	15.2	16.4	16.8	18.6	16.4	18.0
Once in a while	16.7	14.2	22.0	16.2	20.5	15.6
Often	40.0	44.4	40.7	50.8	40.5	49.1
Very often	28.0	25.0	20.5	14.5	22.6	17.3
Total	100.0	100.0	100.0	100.0	100.0	100.0

Note: The number of respondents is contained within parentheses.
*** Data for rural and total are significant at the 0.001 level.

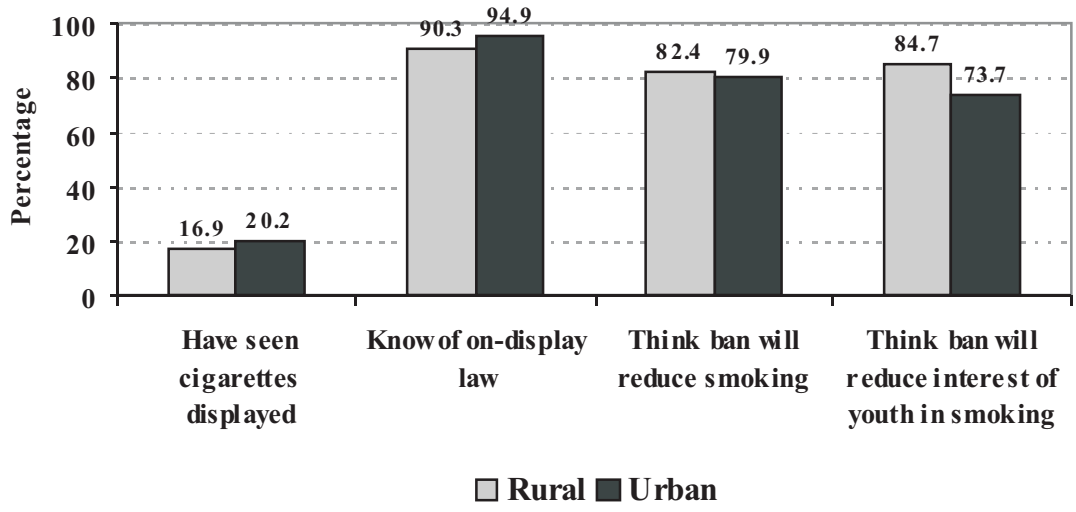
When comparing pictorial versus written warning labels on cigarette packets, over 80 per cent of the smokers in wave 2 reported that the pictorial warnings on the packets were most effective in making people think about the dangers of smoking. Almost all the smokers reported having seen the pictorial warnings (see figure 1), and there was very little difference between the rural and urban areas in the distributions.

Figure 1. Percentage distribution of respondents reporting the most effective type of warning labels on cigarette packets (wave 2)



In the six months before wave 2, less than one fifth of the respondents had seen cigarettes displayed at places of sale and almost all the smokers knew of the law restricting cigarette displays at the points of sale. The majority of smokers agreed that this law would have an effect on reducing smoking and that it would also have the effect of reducing the interest of youth in smoking (see figure 2). Rural respondents were more likely than urban respondents to believe that the law would be effective in reducing smoking and curbing the interest of youth in smoking.

Figure 2. Percentage who had seen displays of cigarettes in the previous six months, knew of the law forbidding the display of cigarettes, thought the law has had an effect and thought that the law has reducedthe interest of youth in smoking, by place of residence (wave 2)



Attempts have been made to reduce the incidence of smoking through increased emphasis on the use of medical consultations to convince people of the dangers of smoking. To some extent it appears that these efforts are working. There was a significant increase in the proportions receiving advice from medical professionals to quit smoking, especially in rural areas (see table 8). However, there was no increase in assistance from medical professionals in helping their patients to quit. There has also been no significant increase in the proportion of smokers receiving information materials at health facilities. The relatively small proportion receiving such materials is disappointing in view of the large amount of information material that has been distributed.

Table 8. Percentage who visited a health professional in previous year who received specified service, by place of residence and survey wave

Service	Urban		Rural		Total	
	Wave 1 (132)	Wave 2 (159)	Wave 1 (393)	Wave 2 (456)	Wave 1 (425)	Wave 2 (615)
Advice to quit smoking	63.6	77.4 ***	66.7	80.3 ***	65.9	79.5 ***
Assistance or referral for quitting	0.0	2.5	3.6	3.1	2.7	2.9
Received pamphlet or other information sheet about quitting smoking	13.5	20.1	20.3	16.9	18.6	17.7

Note: The number of respondents is contained within parentheses.
*** Significant at the 0.001 level.

Among the smokers interviewed in wave 2, the majority had no plan to quit smoking; moreover, there was a significant increase from wave 1 to wave 2 in both rural and urban areas in the percentage who had no plan to quit smoking. One possible reason for this is selectivity: the longer a cohort is followed, the more it will be composed of hardcore smokers. Less that 5 per cent of the smokers planned to quit within the month following the wave 2 survey and a further 16 per cent planned to quit within the next six months (see table 9).

Table 9. Percentage distribution of those planning to quit smoking, by place of residence and survey wave

Plan to quit smoking	Urban		Rural		Total	
	Wave 1 (420)	Wave 2 (347)	Wave 1 (1,138)	Wave 2 (942)	Wave 1 (1,558)	Wave 2 (1,289)
Next month	5.2	3.2	7.7	3.9	7.1	3.7
In next six months	9.5	14.7	15.1	17.0	13.6	16.4
Beyond next six months	21.0	10.1	18.9	12.7	19.4	12.0
No plan to quit	64.3	72.0	58.3	66.3	59.9	67.9
Total	100.0	100.0	100.0	100.0	100.0	100.0

Note: The data for urban, rural and total are significant at the 0.001 level; the number of respondents is contained within parentheses.

In the 18 months between the wave 1 and wave 2 surveys, only slightly over 12 per cent had managed to quit and remain a non-smoker in wave 2, even though over 40 per cent of the smokers in wave 1 reported that they had a plan to quit smoking. While it is true that those who stated that they had plan to quit smoking were more likely to quit smoking than those without such a plan, the majority of those who had a plan to quit were not able to carry through with their plan.

Table 10 shows that health concerns, both the health of the smokers themselves and that of non-smokers who were exposed to their smoking, were the major reasons for smokers to think about quitting smoking. These reasons became more important over the two survey waves. The other main factor that made smokers think about quitting was their desire to be a good model for their children.

Table 10. Percentage who stated that the specified conditions made them think very much about quitting smoking in the previous six months, by place of residence and survey wave

Condition	Urban		Rural		Total	
	Wave 1 (420)	Wave 2 (413)	Wave 1 (1,138)	Wave 2 (1,140)	Wave 1 (1,558)	Wave 2 (1,553)
Personal health	33.6	47.2 ***	37.9	43.2 ***	36.7	44.3 ***
Health of non-smokers	26.7	41.9 ***	34.8	40.9 ***	32.6	41.1 ***
Societal disapproval of smoking	21.2	22.2	27.4	21.0 ***	25.7	21.3 *
Price of cigarettes	17.6	18.0	24.1	19.0 **	22.4	18.7 *
Restrictions on smoking in public places	16.5	19.7	28.0	17.2 ***	24.9	17.9 ***
Information on risks of smoking	17.7	26.0 **	28.7	22.0 ***	25.8	23.1 *
Warning labels on packets	17.3	27.8 ***	27.6	24.8 ***	24.8	25.4 ***
Wanting to be good model for children	39.1	40.2	43.0	34.2 ***	41.9	35.8 ***

Note: The number of respondents is contained within parentheses.

* Significant at the 0.05 level; ** significant at the 0.01 level; *** significant at the 0.001 level.

Although the policies related to warning labels, prices and restrictions on smoking were relatively less important in making smokers think about quitting smoking than were health concerns, one fifth or more of the smokers reported that each of these reasons was important in their thinking about the decision to quit. Furthermore, some of these policies, such as those on warning labels, could be expected to have had an indirect impact by making the smokers more aware of the adverse health impacts of smoking.

In general, smokers in urban areas were increasing likely to have each condition lead them to think about quitting smoking, while in rural areas there was a reduction in the percentage thinking very much about quitting in terms of each condition (except for health).

Discussion and conclusion

The longitudinal data collected by the ITC Project have enabled a close examination of the impact of the tobacco-control policies in Thailand. Although there are several promising signs of success in the attempts by the Government of Thailand to reduce the use of tobacco, the study also indicates some concerns. The study found that the proportion smoking hand-rolled cigarettes had increased, probably as a result of the tax increases on cigarettes, which made factory-produced cigarettes more expensive. Because hand-rolled cigarettes made with local tobacco do not have the associated warning labelling that factory-made cigarettes must display, the shift in the type of tobacco consumption may be reducing the exposure of smokers to information about the health dangers of smoking. This suggests that there is a need to formulate policies and develop regulations that attempt to reduce smoking and raise awareness of the dangers of tobacco use among the smokers who smoke hand-rolled cigarettes. Further rounds of the project will monitor whether the changes observed in smoking behaviour continue.

Most of the sample approved of the prohibition on smoking in hospitals, offices, restaurants and other public places. The proportions agreeing with these bans increased between wave 1 and wave 2. The increasing public acceptance of placing restrictions on smoking provides the Government with the opportunity to aggressively promote tougher anti-smoking regulations.

The policy to place graphic warning labels on cigarette packets has met with success. Knowledge of the health impacts of smoking increased, and most smokers reported that the graphic warning labels were most effective in making them think about the health dangers of smoking. The success of this policy suggests that the Government should increase the number of warning graphics that are used on cigarette packets. However, there is also concern that as smokers, and

indeed non-smokers, become more used to these warnings their impact will diminish. Therefore, it is essential to collect longitudinal data to monitor the impacts of warning labels.

Compared with wave 1, the proportion of smokers in wave 2 who did not have a plan to quit smoking increased. The proportion that had planned to quit within the next six months changed little between the two rounds. These results suggest that there is a need for more vigorous campaigns to influence smokers to reduce or quit smoking.

The study found that only 4 per cent of those who wanted to quit smoking had received assistance or been influenced by public health facilities that are intended to assist smokers to quit. This is a very low proportion in view of the fact that the Government has designated places, such as tobacco-quitting centres, to take responsibility for assisting smokers to stop smoking. The Government should therefore develop and expand the system of helping smokers to quit. It should also disseminate information among smokers about the assistance available to help them quit.

The majority of smokers who had planned to quit smoking did not quit, although those with a plan to quit were more likely than those without such a plan to have quit smoking. These results clearly show that an effective policy to help smokers quit smoking requires two strategies. The first is to convince smokers of the need to stop smoking and the second is to help those who see the need to quit to actually stop smoking.

Convincing smokers of the need to quit is best done by emphasizing the negative impacts of smoking on health, not only their own but also the health of others, especially family members. The new health warning policies appear to be having an impact in this regard. Once a smoker wants to quit, other policies such as the banning of cigarette displays can be effective in reducing the social acceptability of smoking.

How the observed trends will play out in the future will be the focus of the next three years of the current project. Comparison of data from the first two waves with those to be collected in the next three waves will produce greater understanding of the changes in smoking behaviours, attitudes, beliefs and knowledge about the effects of tobacco use. The results of the data collected will enable careful assessment of the immediate and longer-term effects of the tobacco-control policies essential to the formulation of appropriate responses.

Acknowledgments

The data for this analysis are from ITC-SEA (Thailand) surveys conducted in 2005 and 2006 by the Institute for Population and Social Research, Mahidol University. This project is funded by the United States National Institutes of Health and the Canadian Institutes of Health Research and partly funded by the Thai Tobacco Control Research and Knowledge Management Center for the 2006 survey.

References

National Statistical Office (2006). *Cigarette and Alcohol Consumption Behavior Survey 2005,* Bangkok, NSO.

Thompson, M.E. and others (2006). "Methods of the International Tobacco Control (ITC) Country Survey", *Tobacco Control,* No. 15 (Supplement 3).

World Health Organization (1997). *Tobacco or Health: A Global Status Report*, Geneva, WHO.

Managing the Kanchanaburi Demographic Surveillance System: Creation of a Relational Database Management System

Our experience indicates that it is essential when creating any longitudinal database to invest in the development of systems that maintain confidentiality, while affording the basis for the numerous data linkages that are required for longitudinal data analysis.

By Jongjit Rittirong*

Increasing attention is being paid to the collection of longitudinal data. This attention is, in part, a response to the difficulties faced in establishing causal relations with cross-sectional data. However, the collection and use of longitudinal data has brought with it a series of challenges that are not faced by researchers manipulating cross-sectional data. In this article we describe how initial data management models, based on cross-sectional data storage and

* Institute for Population and Social Research, Mahidol Universtity, Salaya, Nakhon Pathom, Thailand, e-mail: prjrt@mahidol.ac.th.

manipulation used in the Kanchanaburi Demographic Surveillance System (KDSS), were found to be inadequate and were replaced by a database system that is consistent with longitudinal data collection, storage and manipulation.

KDSS commenced operations in 2000; the system was designed to monitor population change and to link that change to social, economic and environmental conditions in Kanchanaburi Province, which is located in the western part of Thailand. KDSS is composed of five strata: urban/semi-urban, rice, plantation, upland and mixed economy; each stratum contains 20 villages/census blocks. KDSS covers 80 villages and 20 census blocks in 13 districts of the province (Institute for Population and Social Research, 2001). Three sets of data-collection instruments, namely community, household and individual, are used in the annual enumeration of households and individuals. Spatial data, including the geographical location of each household, are also collected. Approximately 70,000 persons and 40,000 households have been enumerated and recorded in the system.

The initial data administration system was established for KDSS on the basis of "flat files". Such files were maintained separately for each year and for each unit of analysis, namely spatial, community, household and individual data. This database management system was developed on the basis of the long experience of the Institute for Population and Social Research (IPSR) with the collection and analysis of cross-sectional population-based data. The data were stored in SPSS (Statistical Package for the Social Sciences) format, and links among the different units of analysis and over different years were created through matching based on identification numbers.

As soon as the second year of data collection was completed numerous problems were identified in the database system. First, creating the linkages between the various data sets required an excessive input of human resources because of the complexity of the linking process. Second, problems relating to missing data became evident, especially in cases where individuals moved between study areas, and the identification number of individuals and households changed when they moved. Third, data retrieval was often inconsistent when matching data across years.

Moreover, as a member of the International Network for the Continuous Demographic Evaluation of Populations and Their Health in Developing Countries (INDEPTH), KDSS is interested in sharing data with other site members. The major barrier to sharing data is the different data structure of sites. The original data structure of the KDSS data set was a two-dimensional table containing 500-600 numeric variables. Often the data requested from INDEPTH

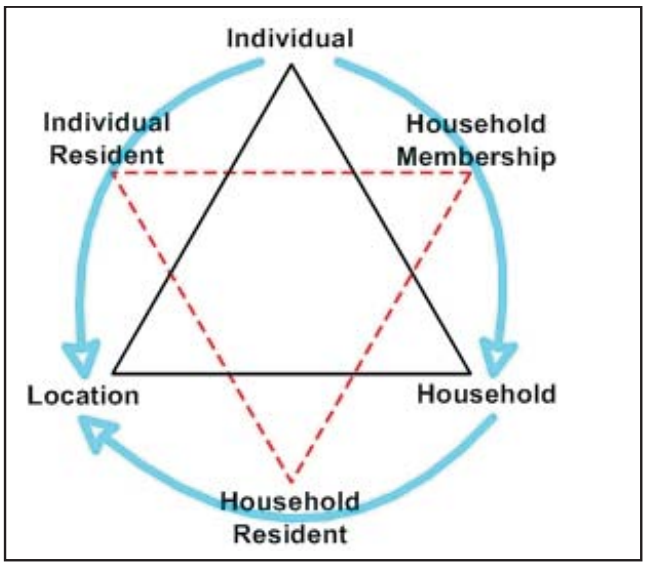
required non-numeric data types, such as date and time data; however, KDSS had split the units day, month and year into three subfields and used a numeric data type for each field. The deficiencies of key fields, such as the start date when individuals appear in the study area and the end date when individuals terminated their stay there, resulted in time-consuming manipulations of the data.

Because of these difficulties, IPSR began the process of changing the database system in order to improve data administration and the reliability of the data. In this article, its experience in transforming the data management system to a relational database management system (RDBMS) is described so that the progress can serve as a lesson to other sites having similar needs.

Relational database management system

KDSS uses the INDEPTH standard approach in order to avoid data redundancy, diminish inconsistency and reduce the waste of resources and to produce a database that is compatible with collaborative database systems (Benzler and Clark, 2000; Benzler, Herbst and MacLeod, 2005). For these reasons, KDSS developed RDBMS, which is operated by SQL (structured English query language) (Lemsiriwongse, 2003). Owing to their concurrent multiple-user accessibility feature, relational databases satisfy many of the information needs of users and make it easier for them to design their own files (Gillenson, 1985). RDBMS is also compatible with spatial data operated on geographical information system software; thus, such data can be integrated into the database system.

Figure 1. Relationship among individuals, households and locations



The transformation to RDBMS was designed to separate the primary formatted data into constant or observation data types and classify related data sets systematically into tables. The related data could then be linked to other tables by created identification information. SQL is an efficient command language that is able to meet conditions and operate calculations simultaneously in order to retrieve data. Output of the SQL operation is compatible with data analysis software, such as SPSS; Stata, R; and SAS (Statistical Analysis Software).

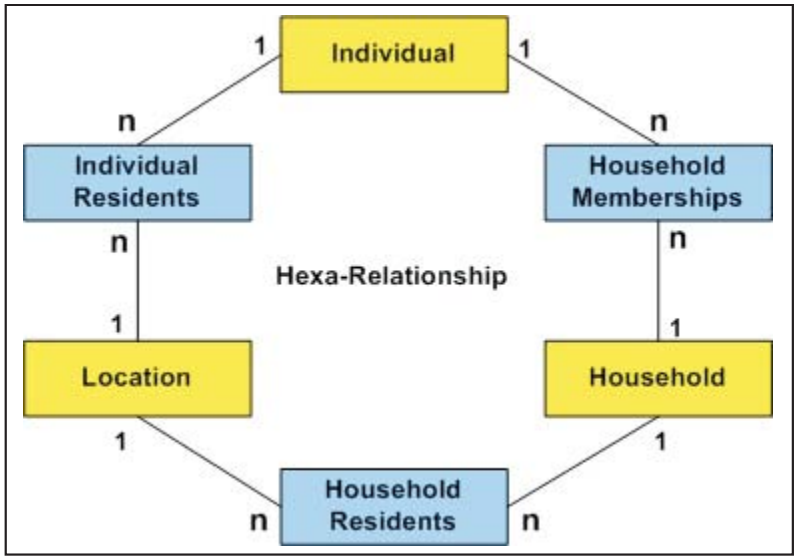
As a member of the INDEPTH Network (INDEPTH Network, 2002), the KDSS database was developed to be compatible with other demographic surveillance system sites on RDBMS (Benzler, Herbst and MacLeod, 2005) according to a demographic data structure called the entity-relationship model, which is depicted in figure 1. The structure includes the relationship among vital events that happen to individuals, households and locations.

A triangle with solid lines links the study objects: individual, household and location. The bottom-up triangle with dashed lines shows the relations or associations among three components: household membership, household resident and individual resident. The objects relate to one another. First, the arrow leading from the individual to the household signifies a relationship between the individual and the household, which is called household membership. Second, the arrow leading from the household to the location represents a relationship between the household and the location, which is called household resident. Third, the arrow leading from the individual to the location denotes a relationship between the individual and the location, which is called individual resident. Cardinalities are the possible numbers of relations; these are shown as a hexa-relationship in figure 2. These relationships simulate the relation of the population factors shown in figure 1. The cardinalities are modelled in RDBMS.

Cardinality (possible relations) can be classified into two types: one (1) and many (n). Figure 2 shows the cardinality between objects, which are described as follows:

- An individual refers to one person
- An individual may be a household member of either one or many households (n households)
- A household may have either one or many household members
- A household may move either once or many times to one or many locations
- A house/resident may contain either one or many households
- A house/resident may have either one or many persons living together
- An individual may or may not change locations

Figure 2. Hexa-relationship



Note: 1 = one and n = many.

The three related objects, individual, household and location, reveal relationships leading to the related vital event database design as well as implementation in the database system. The database structure design covers six necessary data sets shown in the hexa-relationship leading to the theoretical demographic database management system. Moreover, this design allows for future database expansion. The identification connectors, called foreign keys, are recreated, thus providing the means to join the tables. Consequently, the related data can be retrieved through the relation of tables.

Transformation to the new system

The transformation of the KDSS data management system from its original structure to the longitudinal scheme was costly and time-consuming. Table 1 shows the process and the human resources that were required. To accomplish the KDSS transformation to the longitudinal scheme, five steps were undertaken. First, a data investigation and feasibility study on the original data format and an evaluation of data diversity and data volume found that the new system could be operated in conjunction with the original data system. Since the KDSS data set is very specific and requires data in formats that can be used to analyse demographic change, the initial review took six months to determine the methodology for transforming the database. Meetings with the site leader and database staff were undertaken on a regular basis in order to understand researcher needs and potential constraints. Second, because the independently captured data from previous survey rounds needed editing in order to

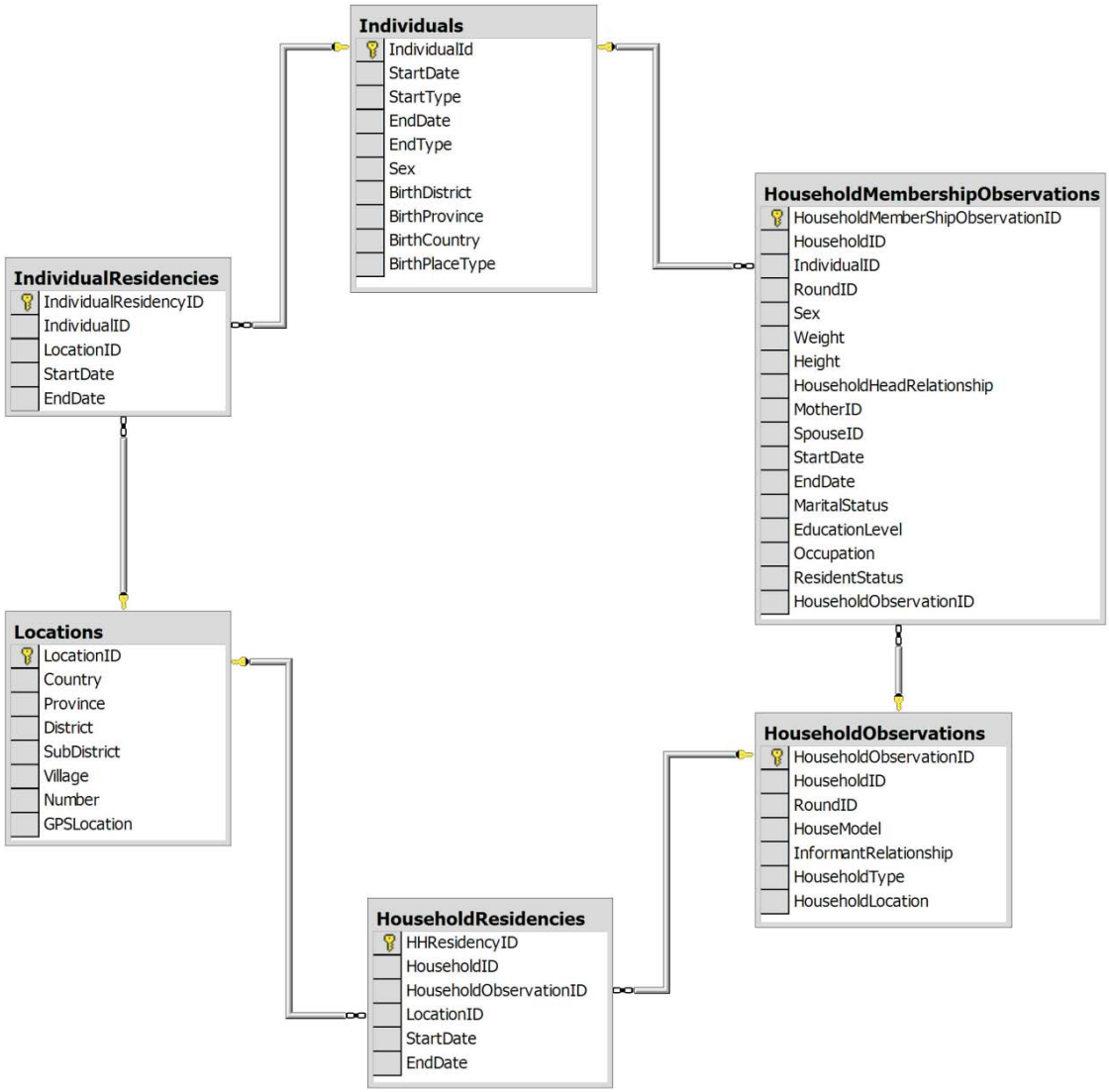
improve data quality and avoid data inconsistency, SQL programming and data correction on the original data set were required to ensure that the cleaned data were input during the transformation process. Third, the database design needed specialized skills that KDSS did not have; therefore, INDEPTH arranged for the Africa Centre for Health and Population Studies, University of KwaZulu-Natal in South Africa to provide technical assistance at the early stages of designing and reviewing the database design. Fourth, when people move to other study areas their individual identification consequently changes; therefore, a new, consistent and confidential identification link was created for use in RDBMS. This step consumed the most time of all, as it was necessary to link individuals over five rounds, which involved approximately 20,000 individuals for each survey round. Fifth, in transferring data to the KDSS relational database, it was necessary to develop an SQL program for automatically transferring the data with specific conditions. The total time for data transformation was 23 months; the process required the work of 20 persons, for a total of 460 person-months, to complete the task.

Table 1. Resources taken to transform the Kanchanaburi Demographic Surveillance System to the longitudinal scheme

Work	Resources		Persons in charge	Activities
	Time (months)	Personnel (persons)		
Data investigation and feasibility study	6	2	Information technology (IT) experts	Meeting at least once a month
Inconsistency checking and data editing	3	7 2	Data clerks IT persons	SQL programming Data correction in the original data set
Database design	1	2	IT persons	Meeting with and assistance from INDEPTH
Linking individuals between years (5-round survey)	9	2 3	IT persons Data clerks	SQL programming
Transfer data to the relational database management system	4	2	IT persons	SQL programming
Total	23	20		

The KDSS data management system consists of three activities: data capture, data retrieval and data updating. As previously discussed, all data in the original database were in SPSS format structured as two dimensions: rows and columns. There were approximately 40,000 records/rows for individuals and 12,000 records/rows for households in each round. Columns contained about 500-600 variables. For the new database, data capture transferred all the data to the new platform employed on RDBMS, which needed a relational data structure, that is, an entity-relationship model.

Figure 3. Conceptual entity-relationship model



Note: The nomenclature in the Kanchanaburi Demographic Surveillance System uses English characters with the first letter of each word capitalized; it avoids special characters and spaces for convenience when querying data by field names.

Individual and household data are included in the conceptual entity-relationship model because they are the data most often used. Nonetheless, community data can also be merged into the KDSS database management system by using a unique identification number on "ID", in this case, the village ID. One of the policies of the project emphasized data confidentiality; confidential information, including name, surname and other means of identification, is not disclosed to any of the researchers. Although the spatial data are bound to the location ID, the geographical individual ID and household ID are masked. Therefore, researchers are able to use the spatial data attached to the individual and household data, but the subjects concerned cannot be identified. To manipulate the new data system, invented IDs created for KDSS are also applied to related research projects. Figure 3 shows the conceptual entity-relationship model for KDSS, which is composed of the following tables:

- *Individuals*: individual information
- *HouseholdMembership Observations*: household membership information enumerated annually contains the relationship between the household's members
- *HouseholdObservations*: household physical characteristics that are enumerated annually
- *HouseholdResidence*: household residency information updated by the most recent enumeration
- *Locations*: global geographic positioning system (GPS) and administration references
- *IndividualResidence*: individual residency information updated by the most recent enumeration

All variables or fields are meaningful in English and each information table is linked by an ID. Figure 3 can be described as follows:

- IndividualID links *Individuals* to *HouseholdMembershipObservations* and *Individuals* to *IndividualResidence*
- HouseholdObservationID, created automatically by the computer system when new records are entered to the database, links *HouseholdMembershipObservations* to *HouseholdObservations*
- HouseholdID links *HouseholdObservations* to *HouseholdMembershipObservations* and *HouseholdObservations* to *HouseholdResidence*
- LocationID links *Location* to *IndividualResidence*

The baseline information, for example the constant individual information contained under *individuals*, is enumerated annually, verified and then recorded once in the database in order to avoid redundancy. Other changeable information is kept as historic information. Each record can be identified by the automatically generated episode ID and date.

The data structure and relationships shown in figure 3 form a basic prototype. The data administrator may design data relationships that are appropriate to meet the requirements of each researcher. An example of the data requests might demonstrate how well users can manipulate the data in the KDSS relational database management system if a data request, adapted from an INDEPTH request, asks for three data tables. First is an individual table containing information on individual ID, sex, marital status, date of birth, start date and end date of the individual's appearance in the study area during the period 2000-2004. Second is in-migration 2000-2004, containing information on individual ID, in-migration date and in-migration duration and on whether or not return migration occurred. Third is out-migration 2000-2004, containing information on individual ID, out-migration date and out-migration duration and on whether or not internal migration occurred. Based on the new system, the data requested can be classified into two groups of variables: constant and observation. The constant variables are individual ID, sex, date of birth and start date. The observation variables are in-migration date, out-migration date, in-migration duration, out-migration duration, end date and information on whether return migration occurred and whether internal migration occurred.

In the original data format, constant variables could be retrieved directly from those fields, except for the start date because it had not been created. However, the retrieval of individuals appearing in the study area at different times might be problematic without duplication. Individuals appear at the first enumeration of the demographic surveillance system survey, while others appear in the ensuing enumeration by birth or in-migration or missed enumeration. The database structure in the new system is designed to store data for distinctive individuals in the *individuals* table, which means that that table contains no redundant data. Observation variables in the new data structure are updated and modified when vital events occur, for example in-migration, out-migration and death. Consequently, event records are updated based on the most recent enumeration. Since an individual might have more than one event record, for instance people might move many times, the database structure can function individually for those events. Out-migration is an example of an event for which each individual could have no record, or one or more records. The out-migration

events are recorded individually and modified by the most recent observation. An out-migration event is terminated when the individual returns to his or her place of origin or dies. The duration of out-migration can be calculated by the function of the difference between the start date and the end date. Information on residency status, including return status, is collected and this is attached to every enumeration. Although those variables are collected separately in tables, they can be linked by a connector called the foreign key. This data manipulation can be done at once with logical conditions and the calculation of the SQL function. This operation responds to the data requested. Our experience has been that the relational database management system for KDSS is less time-consuming than the original database management system, although it does require the skill of an SQL programmer.

The most important lesson learned from the KDSS data management experience is that an identification system should be specified clearly for every unit of analysis and each unit should hold the same identification until the demographic surveillance system is terminated. Therefore, an ID should meet these following qualifications:

- Uniqueness: an ID refers to one individual or one household only
- Confidentiality: an ID contains no direct identification of individuals or households

The KDSS identification system was employed as follows.

Individual identification

(a) *IndividualID* is the ID which researchers use for linking individuals between each year. The ID is unique; it does not contain any meaning that can identify the person concerned. Each individual holds only one ID.

(b) *IndividualExternalID* is the ID that reflects the location meaning. It is used for field work and is not disclosed to the public.

Household identification

(a) *HouseholdID* is the ID which researchers use for linking households between years. The ID is unique and does not contain any meaning that can identify the location of the household concerned. Each household holds only one ID.

(b) *HouseholdExternalID* is the ID that reflects the location meaning. It is used in field work and is not for public access.

Functioning and performance of the system

The database management system can be divided into operations that include data retrieval, documentation and recovery, and data accessibility. Retrieving data from the database can be performed by SQL command. The output data format is compatible with any of the commonly used spreadsheet software. SQL is the standard query language and is similar to other English programming languages; it can be copied and revised according to changing conditions. This reduces the time that must be expended when similar procedures need to be undertaken. In addition, SQL has a feature for linking all designed information from one table to another. Its cascade feature enables users to update and delete all connected tables at once.

Variables, ranges of values, validation system, retrieving data, inserting and updating data can all be managed with the appropriate tools. Data collection and capture should be prepared at the beginning of the project. A mandatory feature of the entire system is flexibility, which is needed to provide accessibility for new users.

Documentation describing meta data or a data dictionary is necessary for researchers who are not familiar with a data set. Meta data or a data dictionary should comprise all information for data utilization, such as variable name description, data type, range of values, period of data collection and conditions. If some values are coded, a code book should be attached using standard codes, such as the *International Statistical Classification of Diseases and Related Health Problems* (10^{th} revision) and the national, provincial/district and subdistrict codes (ICPSR, 2005).

In case a database system "crash", it is extremely important to have a backup and recovery system, and this should be run regularly in and out of house. In addition, a log file containing the updated data history is very useful for rolling back in case there is a database crash.

Data accessibility is based on a set of permissions and a level of authorization to a person for data manipulation in RDBMS. Names, surnames, individual IDs and the household IDs of respondents containing direct means of identification are confidential. Therefore, KDSS keeps these data separately from others and grants data access permission to the data administrator only. Researchers are able to use individual IDs and household IDs but these cannot identify persons or households. Table 2 shows the examples of authorization of levels.

As shown in table 2, only the data administrator is in the administrator group and able to read (retrieve) and write (insert/update) all data. The field manager can only retrieve all data but cannot insert or update any data. The data administrator may set up more groups, such as researcher groups, and assign a permission level to

such groups for accessing specified data tables in RDBMS. New users may be added to the existing groups or newly created groups and assigned a specific level of authorization for tables which they want to use in their research. All users need to be approved by the KDSS project director and/or institute's director before being granted accessibility.

Table 2. Examples of users authorized in the RDBMS

User name	Group	Authorization	
		Read	Write
Data administrator	Administrator	All	All
Field manager	Staff	All	None
Rittirong	Researcher	Specified data table, except confidential data	None

With the transformed database system it is convenient to retrieve data by writing SQL commands which can be revised and re-used. The unique identifiers enable efficient search through all the tables in the database because users are able to set conditions and they do not need to link individual or household annual records for each data retrieval requested. Consequently, this system reduces the time that must be expended and the number of processes required for data retrieval. Data consistency can be verified automatically by specially designed computer programs that run during data capture. As a result, data quality is improved.

Moreover, KDSS operating on RDBMS is flexible and is able to add more records to tables; the systemis also sufficiently flexible for adding values to newly created data description tables. Since the variable and value systems are intended to append new values instead of changing them over various surveys, users do not get confused with the expansion of variables.

RDBMS is efficient for managing a large volume of related data. Although it has no feature for advanced statistical analysis, all data retrieved from RDBMS can be exported to statistical software packages.

Conclusion

Database management plays an important role in KDSS: it provides data from the longitudinal data set that can be analysed and it improves data quality. The operation and access of the initial database system used in KDSS was costly and time-consuming. Therefore, a new system based on a relational database management system was developed to overcome these disadvantages. RDBMS

operates by using structured English query language, which is reliable and sufficiently flexible for operating a longitudinal database. In addition, the KDSS relational database was developed based on the INDEPTH model; therefore, it is compatible for sharing data among other sites in the Network.

To formulate an RDBMS, technical issues must be incorporated within the system: in particular, an identification system should be specified clearly for every unit of analysis. Each unit must hold the same identification until the demographic surveillance system is terminated. Although RDBMS has no advanced statistical analysis functions, it is powerful and able to manipulate the data into formats that are accessible to users. RDBMS is able to update data history, and back up and recover data. These features minimize data damage in case the system crashes.

The experience of IPSR in creating the RDBMS database for KDSS is useful for other research projects that are developing longitudinal database systems. Our experience indicates that it is essential when creating any longitudinal database to invest in the development of systems that maintain confidentiality, while affording the basis for the numerous data linkages that are required for longitudinal data analysis.

References

Benzler, J., K. Herbst and B. MacLeod (2005). *A Data Model of Demographic Surveillance Systems: INDEPTH-Network*, <www.indepth-network.org/publications/indepth_publications.htm, accessed on 5 August 2005.

Benzler, J. and S. Clark (2000). *Longitudinal Database Design,* INDEPTH Workshop, 1999 Annual General Meeting of INDEPTH, Johannesburg.

Gillenson, M.L. (1985). "Trends in data administration", *MIS Quarterly,* December, pp. 317-325.

ICPSR. (2005). *Guide to Social Science Data Preparation and Archiving: Best Practice Throughout the Data Life Cycle,* Ann Arbor, Michigan, United States, the Interuniversity Consortium for Political and Social Research (ICPSR), Institute for Social Research, University of Michigan.

INDEPTH Network (2002). *Population and Health in Developing Countries, Volume I: Population, Health, and Survival at INDEPTH Sites,* Ottawa, International Development Research Centre.

Institute for Population and Social Research (2001). *Report of Baseline Survey Round 1 (2000)*, Salaya, Nakhon Pathom, Thailand, IPSR, Mahidol University.

Lemsiriwongse, O. (2003). *Database Design and Management,* Bangkok, SE-ED Education (in Thai).

Guidelines for contributors

Original contributions are invited, especially papers by authors from or familiar with the Asian and Pacific region. Ideally, such papers would discuss the policy and/or programme implications of population issues and solutions to problems and report on experiences from which others could benefit.

All material submitted for the consideration of the Editorial Board should be in the English language. The paper should not exceed 7,000 words, including tables, figures, references and other material. It should include a short abstract (100 to 200 words) of the issues addressed and the most important findings. The manuscript should be prepared in one of the major word-processing programs and be double-spaced. The margins should be at least 3 cm (roughly 1 in) wide, preferably more for the left-hand margin. If possible, please submit the manuscript as an e-mail attachment to the address given below. If e-mail attachment is not possible, send a hard-copy (a single-sided print copy on A4-sized paper), together with an e-file of the text on CD-ROM or floppy disc.

A complete list of references arranged alphabetically by author should also be included at the end of the manuscript together with a few keywords. Please refer to examples in any issue of the *Journal*. Figures and tables should be supplied separately either as e-mail attachments or in the e-file, preferably in Microsoft® Excel® or any major spreadsheet programme.

Manuscripts are accepted on the understanding that they may be edited. Contributors should submit only material that has not previously been published or submitted for publication elsewhere; and they should so state in their covering letter.

A brief introduction about the author(s)including title and affiliations, should also be submitted.

All manuscripts will be submitted to peer review. The name(s) of the author(s) or other identifying information should therefore be placed only on the title page in order to preserve anonymity.

Manuscripts may be sent by e-mail to the Editor, *APPJ,* at escap-population@un.org; or by airmail post to Editor, *APPJ*, Social Development Division, ESCAP, United Nations Building, Rajadamnern Nok Avenue, Bangkok 10200, Thailand.

United Nations publication
Sales No. E.08.II.F.97
Copyright©United Nations 2008
All rights reserved
Manufactured in Thailand
ISBN: 978-92-1-120-566-4

CONTENTS

Articles:

United Nations
ESCAP

Readers of the *Journal* are invited to suggest topics for future coverage and to comment on the articles already published herein. Correspondence should be addressed to the Editor, Social Policy and Population Section, Social Development Division.

United Nations
Economic and Social Commission for Asia and the Pacific
Social Development Division
United Nations Building, Rajadamnern Nok Avenue
Bangkok 10200, Thailand
Tel: (66 2) 288-1586
Fax: (66 2) 288-1009
Web site: http://www.unescap.org/appj.asp
E-mail: escap-population@un.org

ASIA-PACIFIC POPULATION JOURNAL Vol. 23, No. 3 December 2008 UNITED NATIONS ESCAP