International Journal of Epidemiology, 2015, 423–431
doi: 10.1093/ije/dyu089
Advance Access Publication Date: 11 April 2014
Cohort Profile

Cohort Profile

Cohort Profile: The Malaysian Cohort (TMC) project: a prospective study of non-communicable diseases in a multi-ethnic population

Rahman Jamal,[1]* Syed Zulkifli Syed Zakaria,[2] Mohd Arman Kamaruddin,[1] Nazihah Abd Jalal,[1] Norliza Ismail,[1] Norkhamiwati Mohd Kamil,[1] Noraidatulakma Abdullah,[1] Norhafizah Baharudin,[1] Noor Hamidah Hussin,[3] Hanita Othman,[3] Nor Muhammad Mahadi[4] and the Malaysian Cohort Study Group

[1]UKM Medical Molecular Biology Institute (UMBI), [2]Department of Paediatrics, Faculty of Medicine, [3]Department of Pathology, Faculty of Medicine, Universiti Kebangsaan Malaysia, Jalan Yaacob Latif, Cheras, 56000 Kuala Lumpur, Malaysia and [4]Malaysia Genome Institute, Bangi, Selangor, Malaysia

*Corresponding author. UKM Medical Molecular Biology Institute, Universiti Kebangsaan Malaysia, Jalan Yaacob Latif, Cheras, 56000 Kuala Lumpur, Malaysia. E-mail: rahmanj@ppukm.ukm.edu.my

Accepted 19 March 2014

Abstract

The Malaysian Cohort study was initiated in 2005 by the Malaysian government. The top-down approach to this population-based cohort study ensured the allocation of sufficient funding for the project which aimed to recruit 100 000 individuals aged 35–70 years. Participants were recruited from rural and urban areas as well as from various socioeconomic groups. The main objectives of the study were to identify risk factors, to study gene-environment interaction and to discover biomarkers for the early detection of cancers and other diseases. At recruitment, a questionnaire-based interview was conducted, biophysical measurements were performed and biospecimens were collected, processed and stored. Baseline investigations included fasting blood sugar, fasting lipid profile, renal profile and full blood count. From April 2006 to the end of September 2012 we recruited a total of 106 527participants. The baseline prevalence data showed 16.6% participants with diabetes, 46.5% with hypertension, 44.9% with hypercholesterolaemia and 17.7% with obesity. The follow-up phase commenced in June 2013. This is the most comprehensive and biggest cohort study in Malaysia, and has become a valuable resource for epidemiological and biological research. For information on collaboration and also data access, investigators can contact the project leader at (rahmanj@ppukm.ukm.edu.my).

Key Messages

- This multi-ethnic cohort has provided comparative prevalence rates among the major ethnic groups in Malaysia.
- The prevalence data confirmed the increasing trends of type 2 diabetes, hypertension and hypercholesterolaemia in Malaysia.
- The comparison of the urban and rural populations showed similarity in terms of prevalence of lifestyle diseases due to modernization

Why was the cohort set up?

Malaysia's population of 28.3 million, based on the 2010 national census, is multi-ethnic with the three major ethnic groups making up 95% of the total population.[1] Malays contribute to 63.1% of the population, Chinese 24.6% and Indians 7.3% and the rest is made up of other smaller ethnic groups in East and West Malaysia plus a small population of aborigines.

Non-communicable diseases are fast emerging and becoming the major cause of morbidity and deaths in Malaysia, similar to that in the USA or other developed nations. It is clear that with the increasing modernization and standard of living in Malaysia since its independence in 1957, there has been a major change in lifestyle which includes diet as well as physical activity. The data from the National Health Morbidity Survey II (NHMS II) in 1996 conducted by the Ministry of Health Malaysia showed an 8.3% prevalence of diabetes among the adult population aged $\geq$18 years.[2] The NHMS III in 2006 showed that the prevalence of diabetes has increased to 11.6% and this increased further in the NHMS IV to 15.2%.[3,4] This is rather alarming and there is a similar pattern of increasing diabetes prevalence elsewhere in Asia.[5] For hypertension, the prevalence in the NHMS II was 33% and this increased to 42.6% in the NHMS III. The trend is again similar to some of our Asian neighbours like Thailand and Singapore.[6–8]

As part of the government's increasing efforts to address and investigate the rising trends of non-communicable diseases, the cabinet approved The Malaysian Cohort study in 2005. The top-down approach ensured funding was given to sustain the project at least for the first 5 years. The study proposal was prepared by a team of local experts from various disciplines. Malaysia is a member of The Asia Cohort Consortium whose membership includes South Korea, Japan, China, Taiwan, Singapore, India and the USA.

The Malaysian Cohort aimed to recruit a total of 100 000 individuals from the various ethnic groups. This number is smaller in proportion to the population when compared with the UK Biobank study which has recruited 500 000 participants from a population of 50 million. Nevertheless, we believe it has become a valuable cohort to have, that is now a national resource for researchers in Malaysia as well as providing us with an opportunity to collaborate with international institutions. We have completed the recruitment of the targeted number of participants and we would like to report and share our experience and baseline data with others. As one of the newest cohorts amongst developing nations, we also would like to share the unique experience and the challenges in developing such a study in a tropical and multi-ethnic country like Malaysia.

The primary objectives of TMC project are: (i) to study and determine the roles and interaction of genes, environment and lifestyle in various diseases through a large-scale population cohort study; (ii) to discover biomarkers for cancers and other diseases using the genomics and proteomics approach which would eventually lead to early detection and prevention of diseases; (iii) to consolidate and sustain the initiative for research in life sciences through a systematic discovery programme and also international collaborative research; and (iv) to establish a rich database of information and a bank of biospecimens which will become a national resource for research.

Who is in the cohort?

The Malaysian Cohort study was designed to recruit a total of 100 000 participants aged 35–70 years. The study was approved by the institutional review and ethics board of the Universiti Kebangsaan Malaysia. The study approach included using an interview-based questionnaire and various biophysical measurements plus the collection, processing and storage of biospecimens.

Sampling

The cohort sampling was performed using a mixed approach of voluntary participation (through advertisements and publicity campaigns) as well as cluster and targeted sampling. The cluster sampling was used for the rural areas. The rural areas were chosen from the government's Federal Land Development Authority (FELDA) agricultural scheme which was set up in 1956 and focused on the farming of rubber and oil palm. There are currently about

Table 1. Demographic characteristics of the 106 527 participants in The Malaysian Cohort (2006–12) compared with the general Malaysian population (Census 2010[a])

Demographic	TMC (2006–12)			Malaysian population (Census 2010[a])		
	Place of residence			Place of residence		
	Number of participants	Urban (%)	Rural (%)	Number of people	Urban (%)	Rural (%)
Gender						
Male	44 897	71.8	28.2	14 562 638	70.7	29.3
Female	61 630	71.1	28.9	13 771 497	71.4	28.6
Ethnicity						
Malay	46 782	52.4	47.6	14 191 720	66.6	33.4
Chinese	34 624	96.8	3.2	6 392 636	91.0	9.0
Indian	16 218	86.8	13.2	1 907 827	89.1	10.9
Other	8903	45.0	55.0	5 841 952	54.0	46.0
Age range (years)						
35–44	30 293	80.0	20.0	3 690 093	74.0	26.0
45–54	45 909	70.6	29.4	2 974 602	71.6	28.4
55–64	29 074	64.1	35.9	1 888 618	68.6	31.4
65–70	1251	65.3	34.7	1 427 340	64.4	35.6

[a]see *Population Distribution and Basic Demographic Characteristics 2010* (Department of Statistics Malaysia, 2010[1]).

112 000 settlers working in 103 of these settlements throughout Malaysia, and a total of 75 settlements were sampled. A total of 25 907 invitations were sent out to those who fulfilled the age criteria and 19 467 people (75.1%) responded and were recruited. The FELDA cohort is a relatively non-mobile population and provided an advantage for future follow-up and visits. For the urban areas, the participants were recruited from publicity events which were held in cities, towns, government offices, private agencies and housing areas as well as newspaper advertisements. Between April 2006 and September 2012, a total of 106 527 participants were enrolled into the study. The demographic characteristics of the participants and the comparison with the Malaysian population (as of Census 2010) are shown in Table 1.

The inclusion criteria included being a Malaysian citizen and in possession of a valid identification card, not suffering from any acute illness at the time of study and giving informed consent to the study. Those excluded include those with debilitating illnesses including cancers and those who refused consent. A four-layered written informed consent was taken which covers consent for: (i) the study interview; (ii) the biophysical examination; (iii) blood taking, baseline blood tests and storage of biospecimens; and (iv) future research.

Recruitment centres

The main recruitment centre was based at The Malaysian Cohort office at the Universiti Kebangsaan Malaysia Medical Centre (UKMMC) in Kuala Lumpur. We also had two mobile teams recruiting in the other cities, towns, housing areas and the rural areas. Each of the three recruitment teams consisted of 20–24 personnel including enumerators, phlebotomists, laboratory technicians and data assistants. The mobile teams were also equipped with a mobile laboratory to ensure the preservation of biospecimens in rural areas where electricity supply was a problem, and transportation vans to transport the biospecimens within 24 h from the recruitment sites to the central processing site at the UKMMC. For recruitment in East Malaysia, biospecimens were transported via air shipment.

How are the participants being followed up?

Follow-up and endpoints

Each participant was given a health diary to fill up and return to the TMC office every 6 months. This was to record all illnesses, visits to health facilities, medications and procedures, cost of each treatment and source of payment. Due to the low percentage of the return of these self-report forms, we decided to set up a team to call each participant every 6 months and interviewed them based on the health diary. All participants have either a home phone or a mobile phone. This approach has been successful in getting the follow-up data by phone in 70% of the participants. We have not managed to contact by phone a total of 31 957 (30%) participants, and the reasons for this include not answering the phone (43.2%), voice mail response (21.6%), no ringing tone (15.5%), number not in service

Table 2. Socio-demographic and health differences between those successfully and unsuccessfully followed up by telephone among 106 527 participants of The Malaysian Cohort[a]

Demographic and health differences	Successful (N=74 653)			Unsuccessful (N=31 874)			P-value
	Number of participants	Urban (%)	Rural (%)	Number of participants	Urban (%)	Rural (%)	
Gender							
Male	32 255	74.9	25.1	12 642	64.0	36.0	$\chi^2 = 115.1$
Female	42 398	74.4	25.6	19 232	63.8	36.2	$P < 0.001$
Ethnicity							
Malay	31 603	55.4	44.6	15 179	46.2	53.8	
Chinese	25 764	96.9	3.1	8860	96.2	3.8	$\chi^2 = 1093.9$
Indian	12 026	87.9	12.1	4192	83.5	16.5	$P < 0.001$
Other	5260	51.0	49.0	3643	36.4	63.6	
Risk factors							
Hypertension	34 464	70.7	29.3	14 954	58.8	41.2	
Diabetes mellitus	12 072	68.4	31.6	5523	57.9	42.1	$\chi^2 = 221.5$
High cholesterol	33 106	72.6	27.4	14 014	62.3	37.7	$P < 0.001$
Obesity	12 946	67.9	32.1	5820	57.5	42.5	

[a]All data are row percentages.

(11.8%), wrong number (5.4%) and missing contact number (2.5%). Migration or change of address could also be a cause. The differences between the group which was successfully followed up and the group which we failed to contact are shown in Table 2. A comprehensive follow-up is targeted every 5 years, where each participant will again interviewed, biophysical measurements repeated and biospecimens collected. The invitations to the 5-yearly revisits are issued by phone and through invitation letters posted to their addresses. For the 30% non-responders, our mobile teams will trace them via home visits.

In Malaysia, every citizen is provided with a national identification card (IC) which has a unique number for each individual. For the mortality data, the IC numbers of the participants were sent every 6 months to the National Registration Department (NRD). The NRD provided us with the mortality data and the cause of death.

What has been measured?

Questionnaire and interview

The questionnaire was developed by The Malaysian Cohort Study Group with the assistance of advisers from the Asia Cohort Consortium. Several questionnaires from the Korean Cohort study, the Singapore Chinese Health Study and the Fred Hutchinson Cancer Research Centre, USA, were used with permission as references. The questionnaire covered information on demography, occupational history, use of tobacco and alcohol, diet and physical activity, menstrual and reproductive history (women) and medical history. The diet component consisted of a 24-h recall and a 2-day food record. The physical activity questionnaire was adapted from the short version of the international physical activity questionnaire (IPAQ). The questionnaire was uploaded onto tablet personal computers with touch-screen features. Key pop-up features included a data dictionary as well as a digital diet album, to assist both the enumerators and the participants. Each participant was interviewed face to face at the central recruitment centre or at the mobile sites by a trained interviewer.

Quality control of interviews

We introduced a quality control system for the data obtained from the interviews. Each interview was recorded, with consent, using the tablet computer recording system as well as an MP3 player. Every interview recording was listened to and audited by an independent enumerator. The errors were coded and rectified accordingly.

Biophysical measurements

Each participant had the following measurements taken: height using the Harpenden stadiometer, weight (Seca weighing scale), BMI, waist and hip circumference, waist-to-hip ratio, body composition analysis using the InBody 720 system (Biospace), lung function test using a spirometer (model SP260 by Schiller), blood pressure (HEM-907 model by OMRON) and electrocardiogram. Each measurement was taken three times where possible and

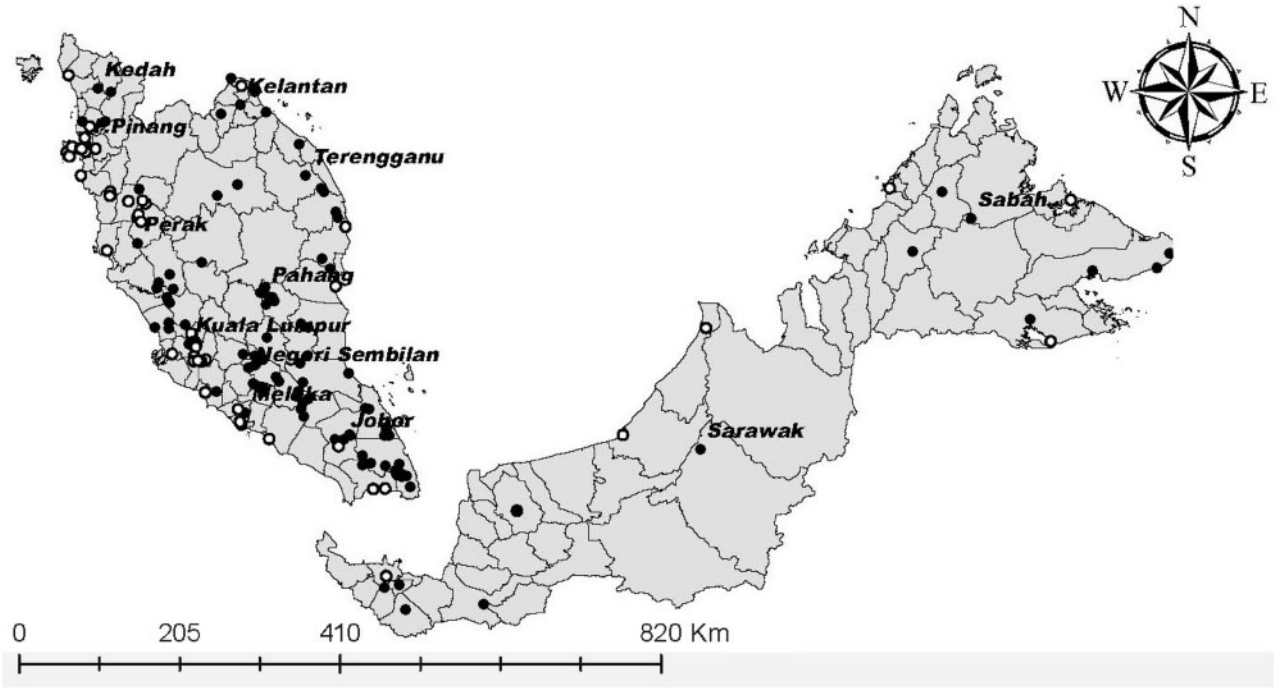

Figure 1. The Malaysian Cohort's 151 recruitment locations, comprising 95 rural (filled circles) and 56 urban (unfilled circles) locations.

underwent quality control processes before the data were uploaded into the database.

Biospecimens

We used the UK Biobank protocol as our reference standards for the collection of blood and urine.[9] All participants came fasted. A total of 40 ml blood and 20 ml urine were donated by each participant. After processing, each participant had 54 cryovials of biospecimens. Half of the cryovials were kept in −80°C freezers and the other half in liquid nitrogen tanks. We used 1-ml Nunc Cryotubes with a 2-D barcoding system (Thermo Scientific, USA) to allow for a systematic inventory system as well as for easy sample tracking and retrieval. The freezers and liquid nitrogen tanks were housed in The Malaysian Cohort Biobank which has an online temperature monitoring system.

Baseline measurements

We measured fasting blood glucose, fasting cholesterol (later full lipid profile after the 50 000th participant), full blood count and renal profile (after the 50 000th participant). Calibration of equipments (Roche Integra and Beckman Coulter) was performed regularly and correlation studies were performed with the hospital's chemical pathology diagnostic laboratory to ensure validity of the results.

Feedback of baseline results to the participants

The results from the biophysical measurements, fasting blood sugar, fasting cholesterol and full blood count were compiled into a one-page summary report and posted to each participant within 2 weeks of recruitment. The report also contained basic explanation on what the normal values were. Those with abnormal results were advised to see their doctor for further investigation and treatment.

Follow-up beginning in June 2013

For the follow-up, we are using the same questionnaire with some minor modifications. Biophysical measurements remain the same but we have added the measurement of cardio-ankle vascular index (CAVI). We are collecting a total of 30 ml blood and 20 ml urine at follow-up. For the blood tests, we added measurement of T4 and thyroxine stimulating hormone (TSH) levels as well as HbA1c levels for those with diabetes.

What has The Malaysian Cohort study found?

Baseline demographic characteristics and habits

As of 30 September 2012, we have recruited a total of 106 527 participants from all over Malaysia. Figure 1 shows

Table 3. Demographic characteristics, educational level, smoking habit, alcohol use and prevalences of diseases according to ethnicity and age group among the 106 527 participants in The Malaysian Cohort (2006–12)[a]

Baseline characteristics ($N = 106\,527$)	Men, by age (years)[a] $N = 44\,897$				Women, by age (years)[a] $N = 61\,630$				Total (%)	Chi-square	
	35–44	45–54	55–64	65–70	35–44	45–54	55–64	65–70		$\chi 2$	*P*-value
Place of residence											
Urban	81.8	73.8	62.0	59.8	78.9	68.4	66.1	73.4	71.4	6.8	<0.001
Rural	18.2	26.2	38.0	40.2	21.1	31.6	33.9	26.6	28.6		
Ethnicity											
Malay	40.2	44.9	50.0	47.6	38.2	46.3	42.7	33.7	43.9		
Chinese	31.4	29.2	28.6	36.3	34.5	32.6	37.6	54.4	32.5	292.4	<0.001
Indian	17.8	17.0	14.6	10.1	16.5	13.6	13.5	7.7	15.2		
Others	10.6	8.9	6.8	6.0	10.8	7.5	6.2	4.2	8.4		
Highest educational level											
University/college	39.5	27.0	20.6	10.3	31.0	20.1	16.9	4.5	24.8		
Secondary school	50.9	51.2	37.0	27.8	56.4	46.9	28.8	21.4	45.6	1173.9	<0.001
Primary school	9.2	20.8	40.3	59.0	11.5	29.3	45.4	53.5	26.6		
No schooling	0.4	1.0	2.1	2.9	1.1	3.7	8.9	20.6	3.0		
Tobacco smoking											
Yes	59.6	58.4	56.6	55.1	5.2	3.2	2.9	4.5	26.6	39 178.0	<0.001
Alcohol drinking											
Yes	11.8	11.4	9.8	9.5	1.8	1.2	1.0	0.8	5.4	4731.5	<0.001
Prevalence											
Hypertension	32.3	46.7	62.1	70.9	25.5	47.4	65.0	80.1	46.5	100.4	<0.001
Malay	30.7	45.3	60.5	72.0	29.6	52.1	69.2	83.6	49.2		
Chinese	31.3	44.6	62.7	68.9	18.6	38.8	58.8	77.2	41.7	608.8	<0.001
Indian	34.1	50.1	63.2	63.5	23.4	45.9	65.3	84.6	45.7		
Others	38.7	54.2	69.9	86.4	35.9	58.7	73.8	81.0	52.9		
Diabetes mellitus	9.9	18.3	26.7	27.3	7.3	15.3	23.3	27.3	16.6	310.3	<0.001
Malay	9.8	18.6	27.0	29.2	8.4	19.1	29.2	32.7	19.2		
Chinese	5.7	10.9	18.6	22.7	3.3	6.7	12.2	22.2	9.1	3376.7	<0.001
Indian	19.9	33.4	45.6	41.9	13.9	25.2	38.0	48.7	28.3		
Others	5.8	12.5	17.5	16.3	5.8	11.8	18.8	9.5	11.1		
High cholesterol	38.7	47.8	51.3	56.5	25.2	45.8	61.6	64.4	44.9	112.1	<0.001
Malay	44.9	54.0	56.6	60.9	30.4	51.5	66.4	66.7	51.0		
Chinese	34.0	42.3	46.3	54.7	21.7	41.4	59.1	63.4	40.8	1381.0	<0.001
Indian	37.8	47.6	49.5	52.7	22.2	40.7	58.6	68.4	41.6		
Others	31.4	35.0	37.9	39.0	22.2	38.8	50.3	52.4	34.4		
Obesity	16.4	14.6	12.1	7.0	19.8	21.3	19.0	10.8	17.7	650.2	<0.001
Malay	18.8	17.7	15.1	10.9	25.7	29.1	26.8	17.5	22.9		
Chinese	11.4	7.8	6.7	1.9	7.8	7.3	7.4	5.1	7.8	3554.8	<0.001
Indian	20.2	18.3	14.0	9.3	29.6	27.9	26.9	23.1	23.4		
Others	15.7	14.5	9.6	4.4	22.6	22.0	18.1	9.5	18.0		
Number of participants	11 451	18 462	14 240	744	18 842	27 447	14 834	507	106 527		

[a]All data are column percentages.

the distribution of all 151 recruitment locations. The breakdown of the participants in terms of age, sex, ethnicity, location (urban or rural), educational level, smoking and alcohol intake is shown in Table 3. There were more women than men. We oversampled the Indians and Chinese to allow us to have more events in these two ethnic groups for comparison with the Malay ethnic group in future research. The prevalences of smoking and alcohol intake among TMC are 26.6% and 5.4%, respectively (Table 3).

Validation studies

Three validation studies are being performed including urine cotinine levels for smoking history, serum

Table 4. Presence of risk factors among the 106 527 participants in the TMC

Risk factors (hypertension, diabetes, hypercholesterolaemia and obesity)	Number of participants	% of total
No risk factor	26 588	25.0
One risk factor		
Hypertension only	11 083	10.4
Diabetes only	1700	1.6
Hypercholesterolaemia only	13 738	12.9
Obesity only	7143	6.7
Sub-total	33 664	31.6
Two risk factors		
Hypertension + diabetes	2112	2.0
Hypertension + hypercholesterolaemia	10 708	10.1
Hypertension + obesity	7519	7.1
Diabetes + hypercholesterolaemia	1875	1.8
Diabetes + obesity	1106	1.0
Hypercholesterolaemia + obesity	4515	4.2
Sub-total	27 835	26.1
Three risk factors		
Hypertension + diabetes + hypercholesterolaemia	3632	3.4
Diabetes + hypercholesterolaemia + obesity	994	0.9
Hypercholesterolaemia + obesity + hypertension	7776	7.3
Obesity + Hypertension + Diabetes only	2304	2.2
Sub-total	14 706	13.8
Four risk factors		
Hypertension + diabetes + hypercholesterolaemia + obesity	3734	3.5
Total	106 527	100.0

carotenoids for fruit and vegetable intake plus a validation study for physical activity using the Actical accelerometer.

Baseline prevalence data and mean values of measurements

The baseline prevalence data of diabetes, hypertension, hypercholesterolaemia and obesity from the 106 527 participants are also shown in Table 3. We used the level of $\geq$7.0 mmol/l as the cut-off point for diabetes [World Health Organization (WHO) criteria] and the 6.21 mmol/l for hypercholesterolaemia (National Institutes of Health, USA).The prevalence of type 2 diabetes of 16.6% is comparable to the 14.6% prevalence obtained from the National Health Morbidity Survey in 2011, although our cohort involved an older starting age group. There are differences in the prevalence of diabetes and obesity between the Chinese (lower prevalence) vs the Malays and Indians. This has provided key opportunities for genome-wide association studies (GWAS) as well as gene-environment-lifestyle comparison between ethnic groups. A genome-wide association study on type 2 diabetes is currently being done. Table 4 shows the presence of risk factors either singly or in combination. A total of 43.4% of the participants have more than one risk factor. The mean values of baseline measurements and blood tests are shown in Table 5. There are differences between the values among men and women and also between the age groups.

Mortality data from 2007–13

Table 6 shows the mortality data and the causes of death since we started recruitment in 2006, up to June 2013. The cause of deaths according to the cancer types are also shown within Table 6.[10] The calculated crude mortality rate for the cohort is 1284 per 100 000 person-years.

What are the main weaknesses and strengths?

Strengths

i. A top-down project approved at the cabinet level ensured the sustained funding from 2005–13.
ii. The establishment of the first large population-based cohort study in Malaysia has comprehensive assessments of exposure, diet and physical activity, biological specimens (blood and urine) and 6-monthly follow-up data.
iii. Many innovative technologies were used, including e-questionnaire (the questionnaire was downloaded to tablet PCs and used by the enumerators to interview the participants), mapping of each participant's address using the Geographical Information System (GIS) and a mobile laboratory for use in the rural communities. The questionnaires were also translated into English, Mandarin and Tamil.
iv. The use of GIS has given us the opportunity to map and layer the environmental data and to facilitate the study of many aspects of diseases including gene-environment interaction.
v. The development of our own in-house Cohort Information Management System (CIMS) manages many key aspects of the study including registration, questionnaire data, biophysical data, results of blood tests, biobank and follow-up data.
vi. Extensive quality control of data includes listening to audio recording of interviews to detect and correct errors, and checking of biophysical data.
vii. The Cohort biobank follows strictly international standards of biobanking and we also use the UK Biobank procedures as a main reference. Our

Table 5. Mean values of baseline measurements and blood tests according to age group and gender among 106 527 participants in The Malaysian Cohort[a]

Mean	Men ($N = 44\,897$) by age group				Women ($N = 61\,630$) by age group				*P*-value
	35–44	45–44	55–64	65–70	35–44	45–44	55–64	65–70	
Systolic blood pressure (mmHg)	124.8 ± 14.5	128.6 ± 16.9	133.2 ± 19.3	137.1 ± 21.3	117.1 ± 15.9	126.5 ± 19.2	133.0 ± 20.5	140.4 ± 22.3	< 0.001
Diastolic blood pressure (mmHg)	82.9 ± 10.9	84.0 ± 11.4	83.1 ± 11.6	81.0 ± 12.0	79.7 ± 11.6	83.0 ± 12.2	82.4 ± 11.6	81.1 ± 11.7	< 0.001
Fasting blood glucose (mmol/l)	5.8 ± 1.7	6.2 ± 2.2	6.5 ± 2.4	6.3 ± 2.0	5.5 ± 1.5	6.0 ± 2.1	6.3 ± 2.4	6.3 ± 2.1	<0.001
Total cholesterol (mmol/l)	5.7 ± 1.1	5.8 ± 1.2	5.9 ± 1.2	5.9 ± 1.2	5.5 ± 0.9	5.9 ± 1.2	6.1 ± 1.2	6.2 ± 1.2	< 0.001
HDL cholesterol (mmol/l)	1.2 ± 0.3	1.2 ± 0.3	1.3 ± 0.3	1.3 ± 0.4	1.5 ± 0.4	1.5 ± 0.4	1.5 ± 0.4	1.6 ± 0.5	< 0.001
LDL cholesterol (mmol/l)	3.7 ± 1.0	3.7 ± 1.0	3.5 ± 1.1	3.6 ± 1.2	3.3 ± 0.9	3.6 ± 1.0	3.7 ± 1.1	3.7 ± 1.1	<0.001
Triglycerides (mmol/l)	1.8 ± 1.3	1.8 ± 1.2	1.7 ± 1.1	1.6 ± 0.9	1.2 ± 0.8	1.4 ± 0.8	1.5 ± 0.9	1.6 ± 0.7	<0.001
BMI (kg/m^2)	26.2 ± 4.4	26.0 ± 4.1	25.6 ± 4.0	24.7 ± 3.6	26.0 ± 5.2	26.5 ± 4.9	26.2 ± 4.8	25.0 ± 4.1	<0.001

[a]Data are means ± SD.

Table 6. Causes of death and number of cases (based on ICD-10) contributing to the mortality in The Malaysian Cohort from commencement of recruitment until June 2013

Main cause of death	Number of cases (%)
Diseases of the circulatory system	440 (32.2)
Neoplasms	
Lung (49 cases)	
Liver (35 cases)	
Breast (35 cases)	
Colorectal (16 cases)	
Stomach (14 cases)	
Nasopharyngeal (13 cases)	
Lymphoma (13 cases)	
Brain (13 cases)	
Ovarian (11 cases)	
Pancreatic (11)	
Unknown (11)	
Other (45)	266 (19.4)
Certain infectious and parasitic diseases	189 (13.8)
Ageing	128 (9.3)
Diseases of the respiratory system	116 (8.5)
Injury, poisoning and similar	55 (4.0)
Endocrine, nutritional and metabolic diseases	48 (3.5)
Unknown/other	31 (2.3)
Diseases of digestive system	29 (2.1)
External causes of morbidity and mortality	24 (1.8)
Diseases of the genitourinary system	20 (1.5)
Symptoms, signs & abnormal clinical and laboratory findings, not elsewhere classified	16 (1.2)
Diseases of the nervous system	4 (0.3)
Diseases of the skin and subcutaneous tissue	2 (0.1)
Total	1368 (100)

bioanalytical laboratory for testing blood sugar, lipid profile, renal profile and full blood count was given the ISO15189 certification from the Department of Standards Malaysia in November 2011.

Weaknesses

i. The The urban Cohort population was somewhat non-representative as we allowed anyone who fulfilled the criteria and those who signed up during our publicity campaigns to contact our call centre and make an appointment to become a participant. However, for the rural community in the agricultural settlements we might have a more representative sample for the population as cluster sampling was used. The data in Table 1 have clearly shown the similarities and differences between the TMC participants and the general Malaysian population. There will certainly be limitations in terms of representativeness and we shall be cautious when using the TMC data in future studies especially in those looking at non-genetic associations.
ii. The failure to contact about 30% of the participants during the 6-monthly phone call was due to migration, transfer of place of work or change in telephone numbers. Measures are being taken to trace them via letters or electronic mail as well as home visits.
iii. Systematic update of exposure data is only now possible as we have completed the target recruitment of 100 000 participants 5 years after the first baseline recruitment. The health diary interview conducted every 6 months covered mainly changes in health status plus treatment.
iv. We had difficulty in obtaining clinical samples at the time of admission for diagnosis of diseases such as cancers. Having tissue samples would certainly add value to future studies.

Can I get hold of the data? Where can I find out more?

Information on The Malaysian Cohort is available at (www.mycohort.gov.com). Requests for other data or

information can be made to the author via e-mail. We welcome national and international collaborations and proposals can be forwarded to (rahmanj@ppukm.ukm.edu.my) and they will then be discussed at the steering committee.

Funding

The project is funded by a top-down grant from the National Biotechnology Division, Ministry of Science, Technology and Innovation (MOSTI), Malaysia (ER-05-01-02-MEB001). The grant is administered by the Malaysia Genome Institute, MOSTI. Additional funding, including infrastructure and utilities, is provided by the Universiti Kebangsaan Malaysia.

Acknowledgements

We thank all the members of The Malaysian Cohort Study Group, all TMC staff members and research assistants. The voluntary participation of all the participants is greatly appreciated. We also thank all the community leaders and the officers from the local authorities involved who have assisted us in the recruitment processes. We are also grateful to the management of the Universiti Kebangsaan Malaysia Medical Centre for providing the space for the TMC office, recruitment centre and the biobank.

The Malaysian Cohort Study Group includes:

Abd Rashid Abd Rahman, Afifah Awang, Ahmad Taufik Jamil, Amal Nasir Mustafa, Amin Ariff Nuruddin, Amin Ismail, Aminuddin Abd Hamid Karim, Amrizal Muhd Nur, Andri Dauni, Atiya Ab Sallam, Christopher Lee, Deepak Rebentisch, Lim Chin Chye, Fauzi Md Anshar, Ghufran Ridzuan, Hashami Bohari, Hayati Aziz, Hejar Abdul Rahman, Ikram Shah Ismail, Ismail Noor, Jamaiyah Haniff, Jamal Hisham Hashim, Jamalludin Ab Rahman, Khadijah Shamsuddin, Krishna Gopal Rampal, Lokman Hakim Sulaiman, Mirnalini Kandiah, Mohamad Ariff Mohd Hashim, Mohd Hashim Omar, Mohd Nizam Isa, Mohd Rizal Abdul Manaf, Muhammad Amir Kamaluddin, Muhamad Nazeri Osman, Nik Mazlan Mamat, Noor Aini Mohd Yusof, Nor Azmi Kamaruddin, Nor Laili Mahat, Norimah Ab. Karim, Osman Ali, Ng Sien Kiat, Poh Bee Koon, Nurmaya Mustafa, Nurul Ain Md Yusuf, Rohaizat Yon, Roslan Harun, Rushidi Ramly, Rusli Ismail, Safiah Mohd Yusof, Salina Abd. Aziz, Salmaan Hussain Inayat Hussain, Salywana A Kamat, Samsuddin Osman Cassim, Saniah Than, Saperi Sulong, Sazaly Abu Bakar, Seri Suniza Sufian, Shahrul Azman Mohd Noah, Sirajuddin Hashim, Suzana Shahar, Syarifah Zainiyah Syed Yahya, Syed Mohd AlJunid, Tan Yoke Hwa, Tengku Mohd Tengku Sembok, Wan Abdul Manan Wan Muda, Wan Ahmad Faisal Wan Salam, Wan Mohamad Wan Bebakar, Wan Nazaimoon Wan Mohamud, Wan Zurinah Wan Ngah, Yasmin Anum Mohd Yusof, Zafar Ahmed, Zainal Ariffin Omar, Zarihah Zain, Zilfalil Alwi, Zulhabri Othman and Zulkifli Ahmad.

Conflict of interest: None declared.

References

1. Department of Statistics Malaysia. *Population Distribution and Basic Demographic Characteristics* 2010. www.statistics.gov.my (18 March 2014, date last accessed).
2. Ministry of Health Malaysia Institute of Public Health. *National Health and Morbidity Survey II*. Kuala Lumpur: Ministry of Health Malaysia Institute of Public Health, 1996.
3. Ministry of Health Malaysia Institute of Public Health. *National Health and Morbidity Survey III*. Kuala Lumpur: Ministry of Health Malaysia Institute of Public Health, 2006.
4. Ministry of Health Malaysia Institute of Public Health. *National Health and Morbidity Survey IV*. Kuala Lumpur: Ministry of Health Malaysia Institute of Public Health, 2011.
5. Ramachandran A, Ma RC, Snehalata C. Diabetes in Asia. *Lancet* 2010;**375**:408–18.
6. Jitnarin N, Kosulwat V, Rojroongwasinkul N, Boonpraderm A, Haddock CK, Poston WS. Prevalence of overweight and obesity in Thai population: results of the National Thai Food Consumption Survey. *Eat Weight Disord* 2011;**16**:242–49.
7. Aekplakorn W, Sangthong R, Kessomboon P *et al*. National Health Examination Survey IV study group. Changes in prevalence, awareness, treatment and control of hypertension in Thai population, 2004-2009: Thai National Health Examination Surveys III-IV. *J Hypertens* 2012;**30**:1734–42.
8. Odegaard AO, Pereira MA, Koh WP *et al*. BMI, all-cause and cause-specific mortality in Chinese Singaporean men and women: the Singapore Chinese Health Study. *PLoS One* 2010;**5**:e14000.
9. UK Biobank Coordinating Centre. *Protocol for a Large-Scale Prospective Epidemiological Resource*. Protocol No: UKBB-PROT-09-06 (Main Phase). Nottingham, UK: Biobank Coordinating Centre, 2007.
10. Jamal R. The Malaysian Cohort Project: Current status and preliminary results. *Symposium 7 Book of Abstracts, Sixth General Assembly of the Asia Pacific Organisation for Cancer Prevention, 26th April, Pullman Hotel, Kuching, Sarawak, Malaysia*, 2012.