DBT supported Bio-repositories, Bio-Bank facilities and Cohort studies

Bio-banking involves the collection, processing, storage, and dissemination of biological samples and their associated clinical data and information, organized in a systematic way. A well-managed biobank is a critical prerequisite for high-quality biomedical research. Recent advances in the tools and technology of molecular biology and genetics have increased the demand for well-annotated, properly preserved specimens. In response to that demand, biobanks have been established on several continents within the past dozen years, and more are in development.

First case study in Denmrk (1993–1997): While investigating the role of oxidative DNA damage as an independent risk factor in cancer way back in 1996, Loft and Poulsen were the first to use the word "**biobank**" **to refer to the use of human biological material** [1]. In a nested case–cohort design, they examined associations between oxidative damage and risk of lung cancer in a Danish population-based prospective cohort of 25 717 men and 27 972 women aged 50–64 years with 3–7 years follow-up. They included 260 cases with lung cancer and a sub-cohort of 263 individuals matched on sex, age and smoking duration for comparison. At enrolment, detailed information on diet, smoking habits, lifestyle, reproduction, medical treatment and other socio-economic characteristics and environmental exposures, including second hand smoke, were collected. Body weight and height were measured and spot urine samples were voided at the clinic and stored at −150°C. Based on the follow up and comparative analysis, they could suggest the association of oxidative damage of DNA with lung cancer in smokers.

Since then, the bio-banking field has grown and improved the conduct of medical research. Much of this progress occurred following the advent of -omics science (genomics, transcriptomics, proteomics, metabolomics) and the ability to develop large electronic databases that store huge amounts of information (big data) associated with patient clinics [2]. In this way, bio-banks have a primary role in the era of precision medicine, which is based on analyzing samples with clinical data. The availability of a large collection of patient samples (with well-annotated patient clinical and pathological data) is a critical requirement for personalized medicine. If more high-quality samples are available through bio-banks, researchers will be able to use these resources to advance patient treatment.

10 Largest Bio-banks in the World:

The following bio-banks are some of the largest in the world, most of which are part of national government funded population-health based studies. There are some notable exceptions such as the Shanghai Zhangjiang Biobank which is industry and government backed. Not all the biobanks have reached their target size. Some biobanks report number of samples rather than participant numbers. A participant may donate multiple samples such as blood, urine, cancer tumors, spinal fluid, fat samples, umbilical cord blood, saliva samples and others such as teeth. When tissues are processed and divided for experiments these are also considered samples and may be stored, therefore biobanks such as "All of Us" biobank will contain considerably more than one million samples when it is completed with a target of 1 million participants. The Shanghai Zhangjiang Biobank has a 10 million sample target, so both "All of Us" and the Shanghai Zhangjiang Biobank are on course to be among the largest in the world, although neither is complete. Biobank Graz has a reported 19.5 million

collected samples, dating back 30 years. Ultimately it is the volume of quality scientific knowledge generated by a given biobank that should be the ultimate measure of its success.

1. **Biobank Graz (Size: 19,546,888 human derived samples):** Biobank Graz is publicly funded and was established in 2007 as a non-profit central research facility of the Medical University of Graz, Austria.
2. **Shanghai Zhangjiang Biobank (Target size: 10 million human derived samples):** A commercial biobank, the Zhangjiang Biobank is located on Li Bing Road in "Zhangjiang Science City", Pudong district, Shanghai, China.
3. **"All of Us" biobank** (**Target size: 1 million participants**): By enrolling one million or more volunteers, the USA based "All of Us" Research Program will have the scale and scope to enable research for a wide range of diseases, both common and rare, as well as increase understanding of healthy states.
4. **The International Agency For Research On Cancer (IARC) Biobank (IBB)** (**Size: samples from 562,000 individuals):** The International Agency for Research on Cancer (IARC) is part of the World Health Organization (WHO). The aim of the IARC is to promote international collaboration in cancer research.
5. **China Kadoorie Biobank** (**Size: over 510,000 participants):** The China Kadoorie Biobank (CKB), known previously as the Kadoorie Study of Chronic Disease in China (KSCDC), is set up to investigate the main genetic and environmental causes of common chronic diseases in the Chinese population.
6. **UK Biobank** (**Size: 500,000 participants):** UK Biobank is a major national and international health resource, and a registered charity in its own right, with the aim of improving the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses – including cancer, heart diseases, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression and forms of dementia.
7. **FINNGEN biobanks** (**Target size: 500,000 participants):** A unique study that combines genome information with digital health care data has been launched in Finland. The FinnGen study plans to analyse up to 500 000 unique blood samples collected by a nation-wide network of Finnish biobanks. The goal is to deepen understanding about the origins of diseases and their treatment.
8. **Canadian Partnership for Tomorrow Project Biobank** (**Size: Over 300,000 participants):** The Canadian Partnership for Tomorrow Project (CPTP) is Canada's largest group of volunteer research participants (population cohort), built to address key questions about what causes cancer and chronic disease.
9. **EuroBioBank network** (**Size: More than 150,000 biological samples):** In 2017, EuroBioBank Network was composed of 25 rare disease biobank members from 9 European countries (France, Germany, Hungary, Italy, Malta, Slovenia, Spain, United Kingdom and Turkey) as well as Israel and Canada.
10. **Qatar Biobank** (**Target size: over 60,000 participants):** Qatar Biobank (QBB), a member of Qatar Foundation for Education, Science and Community Development is a national population-based platform that was initiated to support the future biomedical research and clinical interventions to improve the health of the population of Qatar.

DBT supported Bio-repositories, Bio-banks and Cohort studies:

A. Infectious Diseases:

1. Tuberculosis: RePORT India Initiative - Tuberculosis cohorts and bio-repository (size: about 8000 participants): The Department of Biotechnology is

implementing the Regional Prospective Observational Research on Tuberculosis (RePORT) India Initiative under the auspices of the Indo-US Vaccine Action Programme (VAP), a bilateral programme between Department of Biotechnology (DBT), Government of India and National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), USA.

The RePORT Consortium is a unique research platform, conceived recognizing the necessity for collaborative scientific efforts, within India and with US, to foster Tuberculosis (TB) research and for convergence of diverse groups with similar research focus. The mandate of the RePORT Initiative is to establish a TB consortium with pan India representation, involving long term longitudinal cohorts of TB patients in India. RePORT India was the first consortium to be established as part of the six current global RePORT International networks of TB consortia across the world.

Phase I of RePORT India consortium was initiated in 2013, as part of which, seven Indian institutions collaborated with five Universities from the US to establish cohorts of TB cases and their household contacts. Two major types of prospective observational cohorts of participants are there from whom specimens are collected:

- **Cohort A:** Patients with active TB (including paediatric TB, extra-pulmonary TB and multi-multi-drug resistant TB)
- **Cohort B:** Healthy Household Contacts (HHCs)of TB patients

The total enrollment **of cohort A is 3486 individuals** and **cohort B is 4502 individuals** across all the CRUs. The total number of bio-specimens archived is **127,752 (Cohort A)** and **96,046 (Cohort B).**

These endeavours gain significance in light of the '**End TB Strategy**' announced by World Health Organization (WHO), in 2016, with the target of elimination of TB by 2035 and the Government of India's target, of controlling the TB scourge by 2025. Following the WHO's framework for global TB control, in line with the Sustainable Development Goals of the United Nations (UN), RePORT India is focusing on the 3rd pillar of '**Intensified research and innovation**'. Accordingly, research is in progress to identify biomarkers for relapse and failure in treatment among TB cases, immunology and pharmacokinetics of treatment in pregnant women with TB infection and disease, relationship of TB with the other non-communicable disease epidemic in India, diabetes and treatment response and failure among the MDR TB.

RePORT Sites:

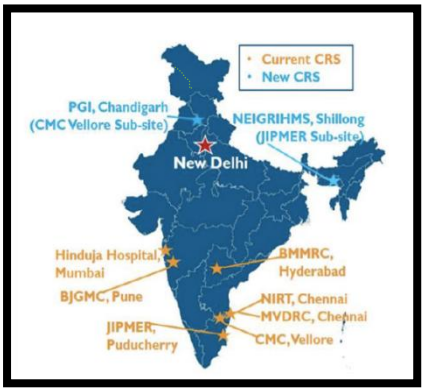


Fig.1: 7 current sites and 2 new sites

The efforts are supported by the establishment of a central bio-repository for specimen storage, located at National Institute of Research in Tuberculosis (NIRT), Chennai and a statistical data management unit at Centre for Health Research and Development (SAS-CHRD), in New Delhi.

2. **Human Immunodeficiency Virus (HIV): Co**horts for HIV **R**esistance and **P**rogression in **I**ndian **C**hildren and **A**dults (CoHRPICA) (Size: HIV-uninfected-1050, HIV-infected without co-morbidity-100 and HIV-infected with co-morbidity-150):

Under DBT-ICMR collaborative effort on HIV/AIDS, a consortia based CoHRPICA Program was initiated in the year 2017. The CoHRPICA Program is the first comprehensive effort towards establishing a consortium of Centers of Excellence with interdisciplinary expertise (clinical, socio-behavioral and biomedical) across India. The Program aims to create uniform standardized cohorts across stages of HIV-infection, the first centralized state-of-the-art biorepository for access to biological samples and a national database to aid in identification of critical research questions and promoting multidisciplinary research collaborations to address the national HIV epidemic.

Specifically,

- The **systematically raised long-term cohorts** will focus on enrolling HIV-uninfected and HIV-infected individuals by accessing relevant populations from the communities (at risk of HIV acquisition) and clinics (providing HIV care) in India. Towards this, the study would leverage on the available capacities and resources at multiple clinical research centers viz. NIE, Chennai; NIRT, Chennai; YRGCARE, Chennai; NARI, Pune and STM, Kolkata which will further link with the National AIDS Control Organization (NACO) through its targeted intervention sites and ART Centers of Excellence spread across the country.
- The **state-of-the-art biorepository** will store and preserve high-quality, clinically annotated biological specimens (blood products and mucosal samples) at the **National AIDS Research Institute (NARI), Pune** to enable their ready access for future scientific research studies. The biorepository will also act as a central resource for storage of samples from other relevant prospective and retrospective studies and resources in India.
- The **central database** (A cloud based server housed at **NIE, Chennai**) will store the demographic, socio-behavioral, clinical and laboratory data collected from the cohorts will link with the biorepository information management system and provide controlled data access to various users through a singular digital platform for epidemiological analyses, generation of new research questions and conduct of advanced immuno-biological analyses.

Together, the above efforts will create an integrated network to enable population-based studies to address the Indian HIV-epidemic by providing context-specific evidence towards understanding the disease and its outcomes and enable design, development and effective implementation of interventions/products/solutions for disease management that are according to the population needs.

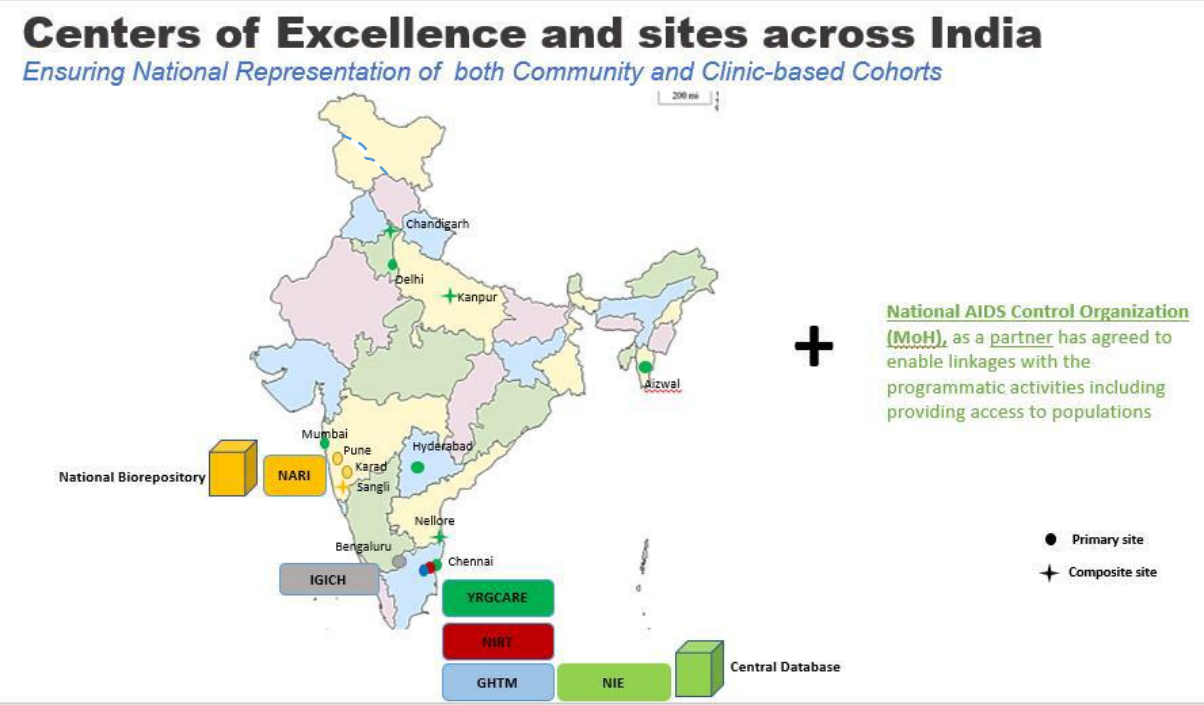


Fig.2: Sites for CoHROICA study

3. **Cambridge-Chennai Centre Partnership on Antimicrobial Resistance in Tuberculosis: Focus on Novel Diagnostics and Therapeutics** (Size: 50 patients with pulmonary MDR-TB and 100 patients with newly diagnosed drug-susceptible PTB)**:**

The need for a **joint India-UK Centre Partnership** focused on drug-resistant tuberculosis (TB) is founded on the scale of the problem both in India and globally, combined with the pressing need to develop new tools and therapeutics to combat it. There were **1,467,585 cases of TB** notified in India during 2012 (the largest number of cases in any country worldwide), with a sharp rise in the number of people diagnosed with multidrug resistant (MDR) TB (from 308 to 16,588 laboratory confirmed cases between 2008 and 2012).

The partnership between the University of Cambridge and the National Institute for Research in Tuberculosis (NIRT) in Chennai, India, will bring together a multidisciplinary team to focus on novel diagnostics and therapeutics for TB. This includes the use of emerging sequence-based diagnostics to improve the accuracy of individual patient treatment for MDR and extensively drug resistant (XDR) TB; new drug targets for TB and prediction/investigation of the impact of resistance mutations based on modelling of bacterial genome data; the development of an in-depth understanding of the bacterial genes in diverse populations of Mycobacterium tuberculosis associated with so-called 'drug tolerance'; and novel approaches to treatment of TB based on immunomodulation (enhancement of autophagy and novel enhancers of T cell responsiveness).

The Centre Partnership will generate a rich and lasting clinical and genomic dataset. Leveraging of novel technologies will increase relevant and collaborative research experience for UK and Indian investigators. Transfer of scientific training and technology to India enhances independent research capacity and fosters future international collaborative projects. This will be achieved through mobility and exchange of researchers.

Five projects:

Project 1: Bacterial genomics as a diagnostic tool in DR-TB;

Project 2: New drug targets for TB through prediction/investigation of impact of resistance mutations;

Project 3: Population based study of gene repertoire associated with drug tolerance and their in vivo expression;

Project 4: Host directed therapy through autophagy stimulation and

Project 5: Manipulating T-cell exhaustion: new therapies to improve outcomes in resistant TB

Clinical Cohorts Methodology: Clinical Cohorts & sample size:

a. COHORT-I - 50 patients with pulmonary MDR-TB (with and without additional drug resistance)

b. COHORT-II - 100 patients with newly diagnosed drug-susceptible PTB (DSTB)

Progress in recruitment and follow-up: Fifty patients in Cohort-I and 100 patients in Cohort II have been recruited. In Cohort-I, 50 patients have completed 6th monthly visit, 30 patients have completed 12th monthly visit and 7 patients have completed 18th monthly visit, respectively. In Cohort-II, 87 patients have completed 12th monthly visit.'

B. Chronic Diseases

I. Neurological Diseases:

1. Indo-Japan collaborated effort: Accelerating the application of Stem cell technology in Human Disease (ASHD): This program is jointly supported by the Department of Biotechnology (DBT) and the Pratiksha Trust, a charitable trust setup by Mr. Kris Gopalakrishnan, co-founder of Infosys and his family. The ASHD program involves top Indian institutes - The National Centre for Biological Sciences (NCBS), Institute for Stem Cell and Regenerative Medicine (inStem), National Institute for Mental Health and Neurosciences (NIMHANS) from Bangalore and the Christian Medical College (CMC) with the Centre for Stem Cell Research (CSCR), a unit of inStem, at Vellore. In order to create and sustain a pool of Indian scientists who can work on stem cell technology, the ASHD program also will partner with the Centre for iPS Cell Research and Application (CiRA), Kyoto University, Japan, which is led by Prof. Shinya Yamanaka, a leading pioneer in stem cell technology.

The ASHD program has two broad components which deal with human diseases of national importance. The first, a collaborative venture between inStem, NCBS and NIMHANS, involves the use of stem cells to study the genetic bases of mental illnesses (such as schizophrenia, bi-polar disease and attention deficit disorders). The second component involves developing modern methods including gene therapy for hereditary blood disorders such as haemophilia, thalassemia and Sickle Cell Disease (SCD), all of which are causes of major morbidity and mortality in India. This component is a joint endeavour between inStem and CMC with the Centre for Stem Cell Research. In order to have maximum impact on these diseases in the community, this initiative also plans to combine these efforts with a community outreach program for the control of major haemoglobin disorders.

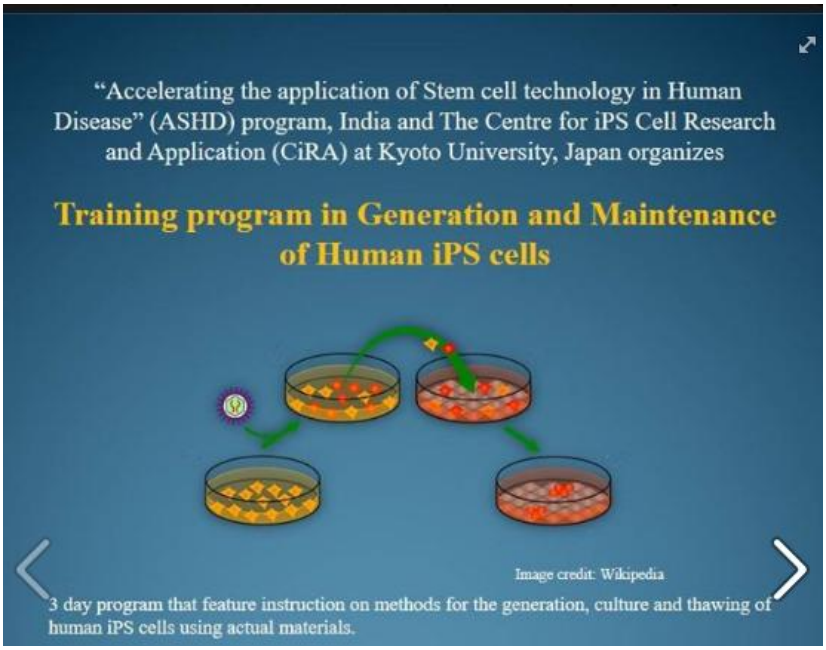


Fig.3: Training program under Accelerating the application of Stem Cell Technology in Human Diseases (ASHD) program

Under this programme following is being supported:

a. Establishing a longitudinal cohort at NIMHANS, Bangalore of patients with 5 mental disorders namely: addiction, bipolar disorders, Alzheimer's dementia, schizophrenia and Obsessive-compulsive disorder (OCD). So far more than 900 recruitments have been done.
b. Establishing a bio-repository of hiPSCs generated from these patients biomaterials at inStem/NCBS Bangalore. So far more than 60 hiPSCs and more than 20 Neural Stem Cell lines have been generated & banked.
c. Establishing bank of iPSC from individuals with most frequent homozygous HLA haplotypes in India at CSCR, CMC Vellore. So far blood samples from more than 200 donors have been collected and are reprogrammed to generate hiPSCs.

2. A Population Based Prospective Cohort Study to Unravel the Causes of Stroke and Cognitive Decline: A Cross-Cultural Perspective (*Sample size* -15000: 7500 urban and 7500 rural population):

Alzheimer's disease, Dementia (unclassified), Frontotemporal dementia, Mild cognitive impairment (MCI), Other NDs not listed, Subjective memory complaints (SMC) or subjective cognitive decline (SCD), Vascular dementia

India is experiencing rapid socioeconomic and epidemiologic transition, and is facing the double burden of pre-transition communicable diseases as well as post-transition non-communicable diseases (NCDs). NCDs largely affect the middle aged and older populations, the groups growing rapidly with aging. The elderly population (60+) is estimated to have crossed the 100 million mark as per the census 2011. The rising numbers of elderly within our population creates a necessity of studying normal aging and cognitive changes associated with the same as this has tremendous implications on providing preventive, and promotive healthcare as well as maintaining the quality of life for elderly people.

The main focus is on investigating subclinical as well as clinical cerebrovascular disease and its relation with cognitive decline in the preclinical stages of dementia. The cohort consists of healthy people above 50 years of age, partly rural/ semi urban and partly urban, both sexes, residents of the selected villages and towns. In particular, no population-based study of determinants of cerebrovascular disease and cognitive decline integrating MRI and genomics has been conducted. Therefore, the study has a potential to yield the information to develop preventive strategies, and to identify future health care needs and policy interventions.

Goals

Overall goal: to establish a large cohort of adults and elderly in the general population as a resource for epidemiological studies of health and aging with a special focus on neuroscience, to generate evidence-based policies for prevention of dementia with aging.

Scientific goal: to study known as well as novel determinants (genetic, socio-economic, lifestyle and environmental factors) of stroke and cognitive function; and to examine unique Indian determinants as well as their consistency and variation of these across the Delhi and Rotterdam cohorts.

Objectives

1. To determine the incidence of stroke and dementia in the community dwelling adults aged 50 years or above.
2. To obtain improved and locally applicable quantitative estimates of importance of major risk factors in term of aetiologic and population-attributable fraction.
3. To define locally valid thresholds for levels of risk associated with various risk factors for myocardial infarction, stroke and dementia
4. To develop risk prediction models for identifying individuals at increased risk of developing stroke (ischaemic or haemorrhagic) or major cognitive disorders, so as to allow personalised prevention strategies.
5. To discover and validate novel biomarkers using recent technological advances in the field of genomics, proteomics, metabolomics, epigenomics and MRI for early (or pre-clinical) detection of major cognitive disorders so as to develop effective tools for dementia prevention.
6. To test and/or develop tools and methods that can serve as models for other Indian researchers/centres to initiate similar cohorts.

Site-1 : Urban (Vasant Kunj & Munirka)

The study completed the target recruitment of 7500 participants at urban site on 23rd July 2019. A total of 7505 participants completed medical site assessments against 8892 who completed field enrollment. Data quality check and cleaning of entered data of all participants has been done. DNA has been extracted for 7325 participants and stored in -80 degree freezers. Funds for genotyping of 7500 DNA samples have been obtained and genotyping is under progress. Six monthly follow up has detected a total of 91 outcome events [Myocardial infarction 47, Death 25, Stroke 19). Regular events committee meetings to ascertain the cause of the events have been organized periodically. Brain MRI has been done for 2584. Feedback from 98% of the study participants have been Excellent or Very good. Three community lectures and meetings with community groups were organized to motivate the people to join the study. A monthly journal club by the research staff of the study has been started to discuss relevant publications.

Site-2 : Rural Site (Ballabgarh)

As on 26th January 2020, a total of 1956 participants have been enrolled, of which medical site assessment of 1614 have been completed. Blood samples of 1646 have been collected and DNA from all samples extracted. Six monthly followup has detected a total of 34 outcome events [Myocardial infarction 20, Death 12, Stroke 2).

New leads obtained

Cross sectional results of baseline assessment has shown a large section (approx. 50%) of sample population who have hypertension but are unaware of it, are untreated or are uncontrolled. Only 30% of those with hypertension are well controlled. Similar situation was found to prevail for risk factors like diabetes, physical inactivity, obesity, hypothyroidism, and Vitamin B12 deficiency. Clearly, there is a huge opportunity to prevent stroke and dementia and also test innovative strategies to improve the control of risk factors. The study plans to test lifestyle interventions like yoga, physical exercise and digital technology to achieve better control of risk factors to prevent stroke and dementia. The planned genome-wide association studies aim to develop prediction models to stratify the population at risk at early stages to promote primordial prevention.

A national workshop entitled ***"Building Knowledge for Prevention of Stroke and Dementia: A National Workshop on Population-Based Cohort Studies"*** at AIIMS, New Delhi on 17th – 18th December 2019 with funding support from DBT and CSIR. Shri Ashwani Kumar Chaubey, Hon'ble Minister for Health and Family Welfare, Govt. of India was the Chief Guest. Several investigators involved in cohort studies from Patna, Mumbai, Kalyani, Bhubaneswar, Raipur and Varanasi attended the workshop.

3. Relationship between Cytochrome P450 G1347A Gene Polymorphism and Risk of Ischemic Stroke in North Indian Population: A Case-Control Study(Size: 250 cases and 250 age and sex matched control subjects):

Stroke is a crucial, devastating neurological disorder which may lead to permanent disability or death. It has been predicted that there will be >200 million Disability Adjusted Life Years (DALYs) lost, 70 million survivors and 20 million deaths after stroke by 2030 across the world [1]. Stroke has been recognized as an intricate, multifactorial and polygenic neurological disorder occurring from the environmental, genetic and combination of vascular risk factors. Heritability presumes the ischemic stroke (IS) varies from 17 to 38% [2,3]. The genome wide association studies (GWAS) have replicated few genetic loci associated with specific IS subtypes. This has guided the concept that genetic risk varies in different subtypes of IS. Several candidate gene association studies have been conducted thus far to identify the relationship of the genetic variants and risk alleles with a particular disease.

In a hospital-based case-control study, 250 cases and 250 age and sex matched control subjects were recruited from Outpatient Department of Neurology, All India Institute of Medical Sciences, New Delhi, India and established the association between CYP4F2 G1347A polymorphism and risk of ischemic stroke (IS) in a North Indian population.
This was an age and sex matched case-control study, undertaken (November 2012 to October 2014) in tertiary referral hospital in India (Neurosciences Centre, All India Institute of

Medical Sciences (AIIMS), New Delhi). A total of 250 patients with 250 controls were enrolled after screening the eligibility criteria.

In this study Cytochrome P450 gene polymorphism is associated with the risk of ischemic stroke mainly for the LVD. This study reveals that CYP4F2 (G1347A) gene variant might be a significant risk factor for IS mainly for LVD subtype of stroke in the North Indian population.

II. Maternal and Child Health

a. Interdisciplinary Group for Advanced Research on Birth Outcomes – A DBT India Initiative (GARBH-Ini) (Size: 4500 individuals):

Globally, preterm birth is a major public health problem. In India, 3.6 million of the 27 million infants born annually are preterm. Risk stratification of women based on multidimensional risk factors assessed during pregnancy is critical for prevention of preterm birth.
GARBH-Ini is a pregnancy cohort to study multidimensional correlates of preterm birth in India. A unique collaborative interdisciplinary program; coordinated by THSTI in partnership with National Institute of Biomedical Genomics (NIBMG), Kalyani; Regional Centre for Biotechnology (RCB), Faridabad and district Gurugram Civil Hospital (GCH), Gurugram and tertiary care hospitals (Safdarjung Hospital, Maulana Azad Medical College (MAMC), New Delhi).

The cohort (GARBH-INI cohort) of pregnant women was started in May 2015 at the civil hospital in Gurugram, Haryana, India with the objectives to identify the clinical, epidemiological, genomic, epigenomic, proteomic and microbial correlates, discover molecular risk-markers by using an integrative omics approach, and generate a risk-prediction algorithm for preterm birth.

Women are enrolled within 20 weeks of gestation and are followed until delivery and once postpartum. The objectives are to identify clinical, epidemiologic, genomic, epigenomic, proteomic, and microbial correlates; discover molecular-risk markers by using an integrative -omics approach; and generate a risk-prediction algorithm for preterm birth. They describe here the longitudinal study design, methodology of data collection, and the repositories of data, biospecimens, and ultrasound images being created.

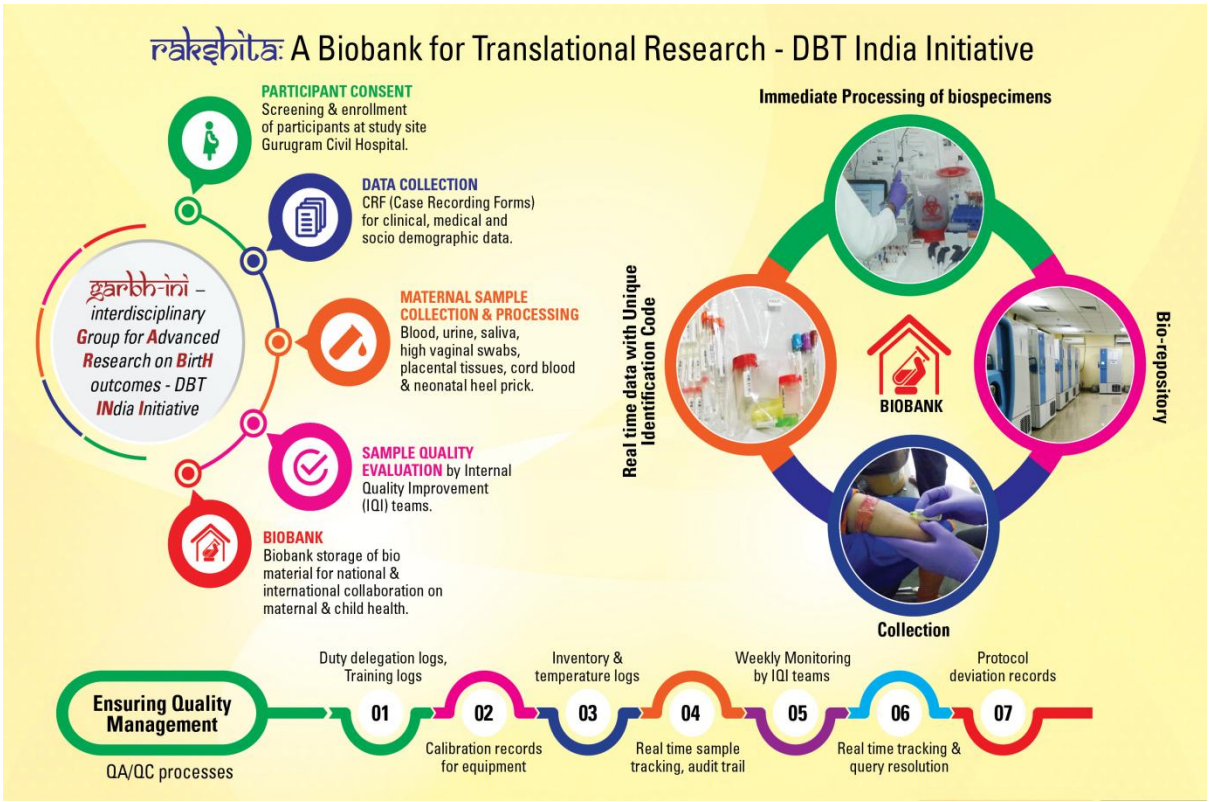


Fig.4: Work plan and strategy of Garbh-ini program

Fig.5: Biobank for translational research under Garbh-ini program

A total of 4,326 pregnant women, with documented evidence of recruitment before 20 weeks of gestation, have been enrolled through March 2018. They report baseline characteristics and outcomes of the first 2,000 enrolled participants. A high frequency of preterm births (14.9%

among 1,662 live births) is noteworthy. The cohort database and the repositories will become global resources to answer critical questions on preterm birth and other birth outcomes.

b. The Pune Rural Intervention in Young Adolescents (PRIYA) study (Size:

The Pune Maternal Nutrition Study (PMNS) was established to prospectively study the relationship of maternal nutrition to fetal growth and later cardiometabolic risk in the offspring. The cohort was set up prospectively in 1994 in six villages near Pune city. At that time, they identified 2,675 married non-pregnant women in 6 villages near Pune for possible enrolment in the study, of whom 2,466 consented to take part. At that time, people's main livelihood was subsistence agriculture. Since the area was drought-prone and lacked irrigation, malnutrition was common. The women recruited to the PMNS (F0 generation) were undernourished and did heavy farming work even when pregnant. Their children (the PMNS Cohort, F1 generation) have been followed up continuously.

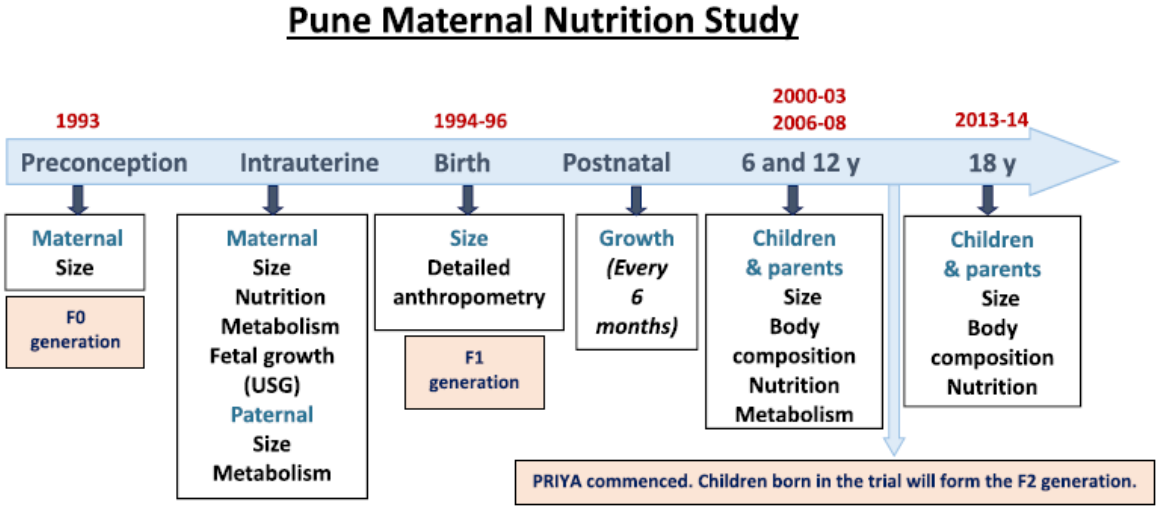


Fig. (Measurements and data available at each stage in the Pune Maternal Nutrition Study). A biobank of stored plasma, serum, DNA, urine and stool samples are available from parents (F0) and offspring (F1) at various time points. Cord blood and placenta samples from the children born in the trial (F2) are collected for genetics and epigenetics, immunophenotyping, and microbiota

Findings from the Pune Maternal Nutrition Study:

a. Maternal diet and micronutrient status, and physical workload, during pregnancy were related to the size of the newborns.
b. Newborns had a 'thin-fat' phenotype (low overall weight and muscle mass but relatively high fat mass) compared with Europid babies, which persisted through childhood.
c. Many of the PMNS mothers had low plasma vitamin B12 levels (70%). Further studies have shown that this is mainly due to low dietary intakes and not to malabsorption. Folate deficiency was rare. Low maternal vitamin B12 status and high folate status during pregnancy predicted higher insulin resistance in the children.
d. At the last completed round of cardiometabolic risk marker measurements (aged 18 years), 28% of the young men and women had abnormal glucose tolerance (either

impaired glucose tolerance of impaired fasting glucose). Glucose concentrations were higher in individuals whose birth weight had been lower. There was tracking of glucose values from early childhood.

The objective of the original cohort study was to explore associations of maternal diet, physical activity and nutritional status with birth outcomes and the evolution of cardiometabolic risk in the children. Since 2012, when the cohort was aged ~18 years, this cohort study has become a randomised controlled trial, testing the effect of vitamin B12 supplementation of the young F1 men and women on fetal growth, epigenetic marks and cardiometabolic outcomes in the next (F2) generation.

In the initial study, women's pre-pregnancy anthropometry was measured, and serial data on their diet and micronutrient status were collected during pregnancy. Fetal growth was measured by ultrasound and detailed anthropometry was carried out at birth and annually during childhood (N=702). At 6, 12, 18 and 24 years, body composition, cardiometabolic risk factors and cognitive function have been measured in the children and parents. One of the main early findings of PMNS was that 70% of the F0 mothers were vitamin B12 deficient during pregnancy, and that this was associated with increased insulin resistance in the F1 children. This, and other evidence, led to the decision to convert the PMNS cohort into an intervention study ("PRIYA") in 2012.

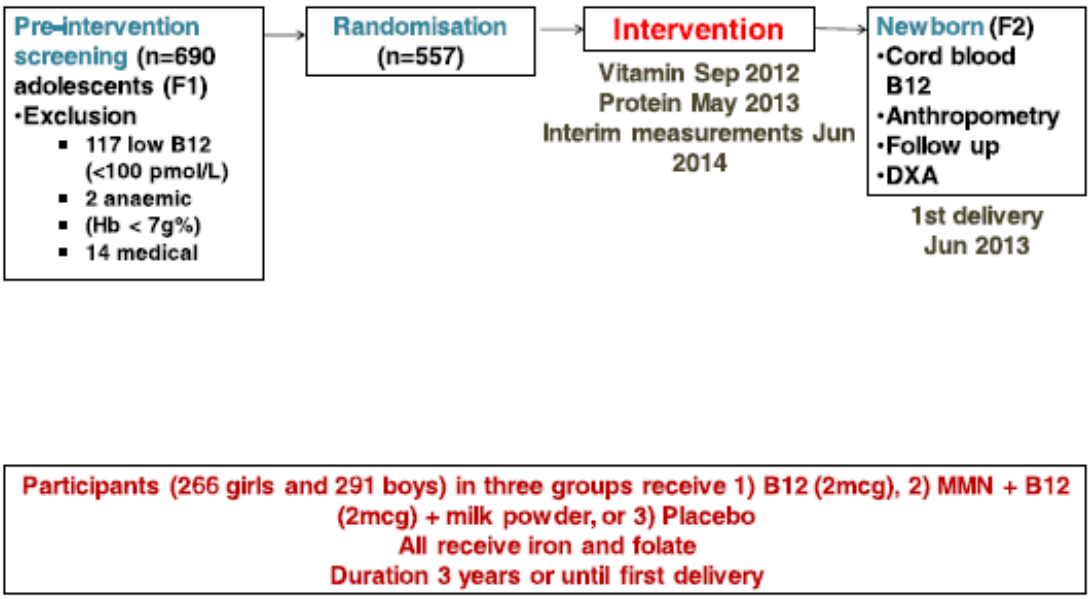


Fig. (**The design of the Pune Rural Intervention in Young Adults study**). The design is a randomised placebo controlled trial with three arms Participants were randomised individually into one of three groups to receive either: i) 2 µg B12, ii) 2 µg B12 plus multiple micronutrients (MMN) plus 20 g milk powder (equivalent to 5 g milk protein), or iii) placebo daily.

III. Renal Diseases

1. Pediatric Renal Biology Program Research on Nephrotic Syndrome, St. John's Research Institute, Bengaluru (Size: 300 individuals):

Nephrotic syndrome (NS) is the most common glomerular disease seen in the pediatric age group. It is the second most common kidney disease seen in pediatric nephrology clinic, the most common being congenital malformations of the kidney and the urinary tract. It is characterized by massive proteinuria, hypoalbuminemia, peripheral edema and hyperlipidemia. This is a multicentric collaboration for research on nephrotic syndrome, the most common cause of chronic renal disease in children. Development of a disease registry shall allow sharing of pre-determined variables on tow longitudinal cohorts of patients with nephrotic syndrome and form the framework on which inter-institutional clinical studies can be subsequently conducted. Another important aim of the collaboration shall be to set up a biorepository at participating institutions with common standard operating procedures. Data entry for both the registry and biorepository shall be anonymized so as to protect patient identity and maintain confidentiality.

Information from the registry and materials obtained through the biorepository shall be available to collaborators across partnering institutes to address questions on disease pathogenesis and therapies in domains selected by them on the basis of their relevance and available expertise. Emphasis on clinical and bench research shall result in improved understanding of mechanisms of disease and precise phenotype- genotype correlation, especially in patients with steroid resistant nephrotic syndrome. On the medium term, this will enable research on novel and focused treatment options that impact the course of the illness, transforming into better clinical outcomes. On the long-term, this nationwide collaboration will result in capacity building in research, and foster opportunities for education and training for postgraduate students and younger faculty.

2. Indian Chronic Kidney Disease Study (Size: 4000 patients):

The Indian Chronic Kidney Disease (ICKD) study is the first kidney disease network in India and has established the largest prospective cohort of patients with CKD in developing countries. The study has established subject enrolment and follows up protocols at 11 centres across India. It is coupled to serial annual biobanking. The electronic records of sample storage and location as well as SOPs for biobanking have been established.

The ICKD study was set up in a phased manner. The planning phase included securing regulatory approvals at all participating centers, appointment of research staff, purchase of equipment and consumables and development of a secure ICKD study database for clinical and lab management. This was followed by a short pilot phase at each center where study procedures were tested and established. The final phase of enrolment and follow up is continuing.

Its success is illustrated from the facts that it has been able to enroll approximately 4000 patients and is a part of the International Society of Nephrology's (ISN) global network of cohort studies in CKD i.e. iNET-CKD (International Network of Chronic Kidney Disease cohort studies).The ICKD study has established a serial annual bio-bank of biological specimens (blood and urine) at a central facility (PGIMER) as well as at each center. The bio-bank is one of its kind in the country and has been digitized with respect to location, tracking and transport of samples. In fact, the ICKD study team is coordinating with investigators of other multi-centric studies in the country to help them develop similar digital solutions.

The ICKD study team is continuously innovating based on the feedback from investigators, participants and audit of data. In order to ensure completeness of follow up data, the study team has designed a private and secure mobile/tablet-based application that will allow exchange of information between study coordinators and participants. The application has been developed and is currently in the testing phase.

IV. Liver Diseases

1. National Liver Disease Biobank (NLDB) (Size: 5.4 million bio-samples):

The National Liver Disease Biobank (NLDB) is a DBT-sponsored facility established at ILBS, New Delhi (http://www.nldb.in) to accelerate deliverable basic and translational research in the field of Hepatitis C (HCV), acute and chronic Liver and Biliary diseases, Gallbladder and various hepato-biliary tumors. It is the first Liver disease biobank in India to provide researchers and industries with high quality biosamples and the patient data with follow-up in order to facilitate high quality research in the field of liver disease, genetics, biomarker research, molecular diagnostics, drug discovery, and new therapeutics.

Fig.6: Cry-storage facility at ILBS under NLDB program

Fig.7: Analytical facility at ILBS under NLDB program

V. Cancer

Cancer Bio-bank: ACTREC, Mumbai (Size: 10,00,000 samples):

ACTREC Biorepository facility is operational since 2007. ACTREC biorepository is the facility to collect, annotate, store, and distribute biological samples to in house researchers under a specified mechanism for duly approved research projects. The Bio specimens are collected from Operation Theatre, Frozen room and Surgical Pathology as well as breast OPD. After taking the ethical consent from the patients they accrue tissue samples from patients taking care not to compromise the diagnosis. Majority of the samples were Head and Neck tumors, followed by Breast tumors. Other tumor types included Gastrointestinal, Neurological and others. For all the possible cases paired normal samples were also collected and stored. The samples are annotated and cryo-preserved and distributed to IRB approved projects as per request.

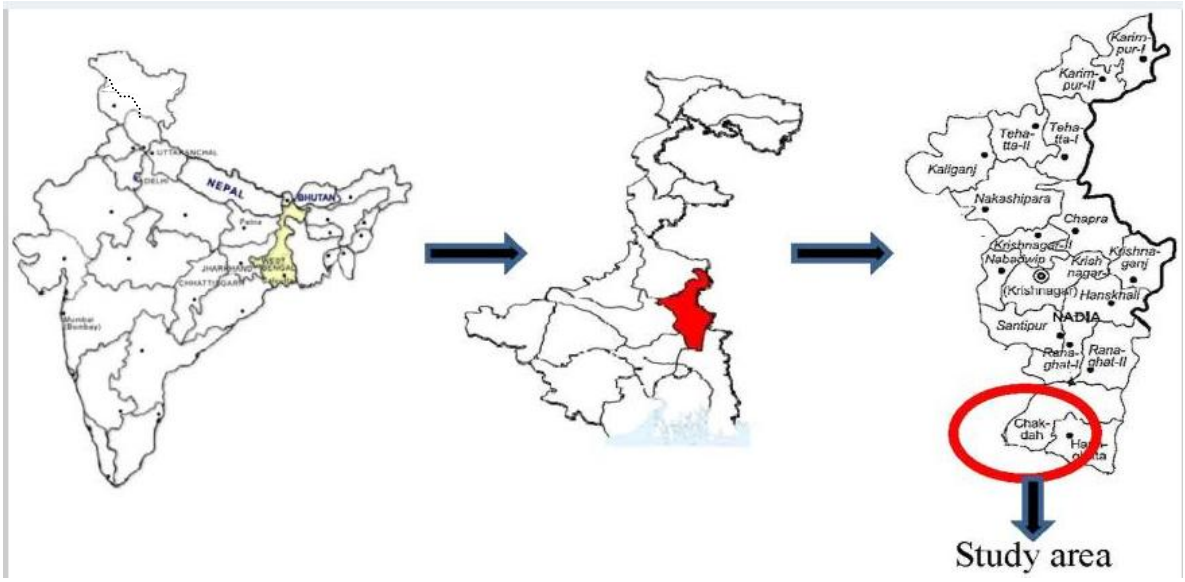
The DBT funded International Cancer Genome Consortium -India project was approved and initiated in 2009 and since then they have been accruing the oral Gingivo buccal squamous cell carcinoma patients under this project.

A pioneer in the field, it is the first tissue bank in India to use radiation for the sterilisation of biological tissues, and in 2004 became India's only Tissue Bank with an ISO 9001:2000 Certified Quality Management System. Currently it banks gamma-irradiate, human amnion (the inner membrane of the sac that encloses the foetus in the womb), skin and bone, and promotes and co-ordinates the development of tissue banks using radiation sterilization of grafts.

VI. Genetic predisposition

A platform for research on complex traits and diseases, Kalyani (Size: 20,000 individuals):

This cohort of about 20,000 individuals drawn from villages in and around Kalyani – a peri-urban region near Kolkata (formerly Calcutta) has been formed to serve as a platform for prospective studies on genomics of health and disease, to be carried out under the leadership of the National Institute of Biomedical Genomics (NIBMG), Kalyani.



The Kalyani cohort created in 2010 by the National Institute of Biomedical Genomics, West Bengal, India, is designed to serve as a platform for conducting prospective basic and translational studies on epidemiology and genomics of health and disease-related parameters, particularly of non-communicable diseases (NCDs). The overall goal is to assess behavioural, biological, genetic, social and environmental factors and obtain necessary evidence for effective health improvement. Collected baseline data comprise 15727 individuals, >14 years of age from seven municipal wards in the Kalyani and Gayeshpur regions. Data are being collected on demographics, current health status, medical history and health-related behaviours. Blood samples were also collected from a subset of individuals (n = 5132) and analysed for estimation of known markers of NCDs. DNA has been extracted from blood samples and stored for future use.

Important baseline findings include a high prevalence of diabetes, dyslipidemias and hypothyroidism. Prevalence estimates for these disorders obtained from self-reported data are significantly lower, indicating that participants are unaware of their health problems. The identification of 'at risk' individuals will allow formation of sub-cohorts for further investigations of epidemiological and genetic risk factors for NCDs. Access to the resource, including data and blood samples, created by this study will be provided to other researchers.

VII. Obesity in children

Healthy Life Trajectories Initiative (HeLTI) (size: 400,000):

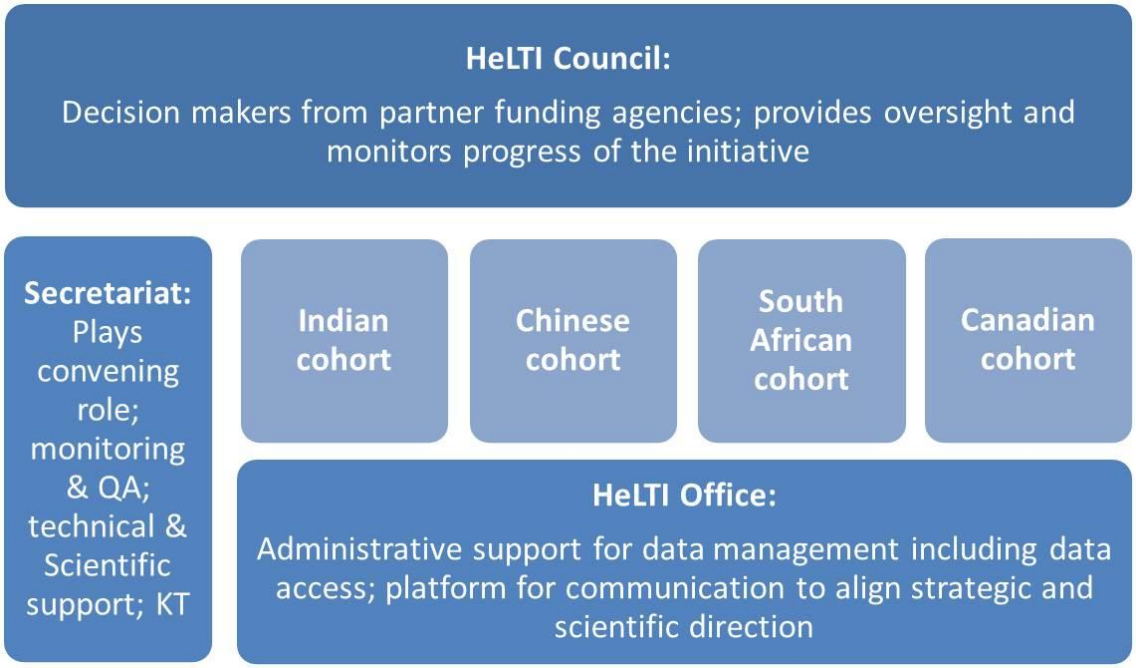
The Healthy Life Trajectories Initiative (HeLTI) was launched at the Developmental Origins of Health and Disease (DOHaD) 10th World Congress in Rotterdam, The Netherlands on 16th October 2017. Globally, the prevalence of obesity in children is increasing and reflects changing patterns of behaviour towards more unhealthy diets and physical inactivity. This is true in high income countries such as Canada and also in countries experiencing economic growth and demographic transition such as China, India and South Africa.

Recognizing this need, the national research funding agencies of Canada, China, India and South Africa, in collaboration with the World Health Organization, have established a programme of research to generate evidence that will inform national policy and decision-making. The collaboration is known as the Healthy Life Trajectories Initiative (HeLTI).

Intervention cohorts will be established in Canada, China, India and South Africa that will examine the effects of interventions starting preconception, during pregnancy and into early childhood on body composition of children at 5 years of age and measures of early child development.

To date, joint funding mechanisms have been established between CIHR and the national research agencies in each country; research teams have been selected based on their knowledge of the science, experience to implement complex intervention studies and ability to sustain cohorts; a joint proposal development workshop has been held to develop a common intervention framework and agree core outcome and process measures.

The following diagram illustrates the design of the Linked International Intervention Cohorts (LIIC):



HeLTI-India Study

Similar to the other HeLTI studies, The HeLTI-India study builds on the Developmental Origins of Health and Disease (DOHaD) concepts to implement a multifaceted integrated intervention programme starting pre-conceptionally and continuing through pregnancy and infancy to enable at-risk children to reach their full potential for health and well-being. It is anticipated that the intervention will reduce childhood adiposity, improve cardiovascular and metabolic health, and improve child development outcomes.

The intervention comprises measures to improve maternal physical and mental health, nurturing skills, infant infection prevention, and to improve child growth and development in the short term and reduce the burden of non-communicable disease and optimize human potential in the long term. (Details of the intervention are described below). The primary outcome at age 5 years in the children is adiposity, measured by fat mass index. Other key outcomes at 5 years of age include overweight and obesity (OWO), glucose metabolism, blood pressure, and infant/child development.

The study is a community-based, cluster randomized intervention with three arms (pre-conception, pregnancy and control) set in rural Mysore, South India, with individual villages forming the basis for the cluster. The team will assess the efficacy of the longitudinal intervention and evaluate the cost-effectiveness of this package. Comparisons of relative benefits and costs will be made starting the intervention pre-conceptually versus initiating it during pregnancy. Recruitment will target married women 18 years of age or older. There will also be local community involvement from fathers and other family members.

Progress

Formative work to help inform and implement the main study has commenced. Local research staff have conducted an enumeration survey to identify eligible women. Community engagement activities were completed within 3 of the study villages. Perspectives were gained from the staff in their role of supporting change and from community leaders, village women and their families about the study.

As shown below during an initial engagement session, the study was well received and numerous women expressed interest in participating. The two main factors contributing to their decision were the potential to benefit future generations; and the possibility of scaling up the intervention to the state or national level.

HeLTI-India Intervention

Pre-conception Interventions (underpinned by behaviour change)

a. Nutrition – micronutrient supplements; diet diversity
b. Avoidance of toxic exposure
c. Hygiene promotion

d. Deworming

Pregnancy Interventions

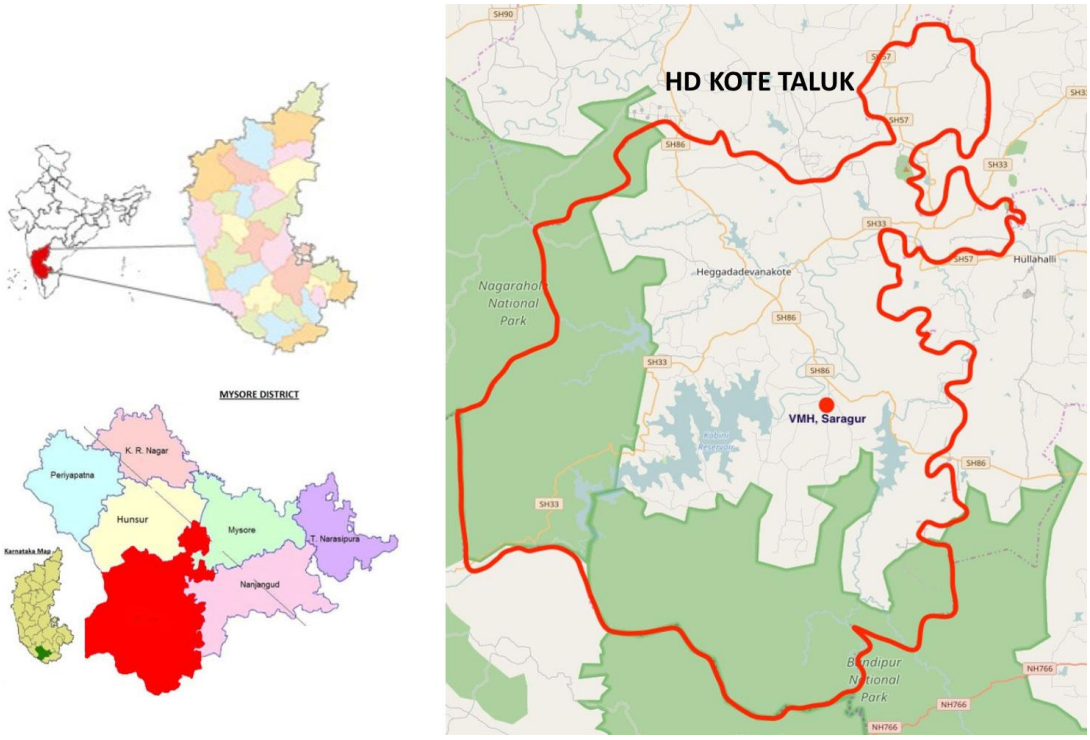
a. Preconception +
b. Group parenting with a cognitive behavior therapy program
c. Birth preparedness
d. Breastfeeding preparedness

Postnatal Interventions

a. Nutrition (maternal)
b. Nutrition (child)
c. Vaccinations
d. Continued group parenting/cognitive behavior therapy

Setting

The main site of the study is Saragur, rural Mysore with a population of **400,000**. The majority have poor nutritional status and limited access to specialized health services.



TEAM INDIA AND COLLABORATORS

VIII. Systemic Lupus Erythematosus

SLE cohort SGPGI, Lucknow (Size: 2500 individuals):

About 20% of systemic lupus erythematosus (SLE) starts in childhood and children have less gender bias in favor of females as compared to adults. Systemic manifestations, nephritis, neuro-psychiatric disease and cytopenias are more common in children at presentation than adults. Since most children develop lupus in their early adolescence, dealing with the diagnosis of an unpredictable lifelong disease during this phase of life is challenging. Physicians must recognise specific medical and social needs of this age group, for optimal long-term outcome. Steroids and immunosuppressive drugs are the cornerstone for treatment in children as with adults with lupus. The out-come has improved considerably with these drugs and 10-year survival is nearly 90%. Due to longer life spans more damage accrues in children as compared to adults. Most of the drugs are associated with significant toxicity and the goal of having a drug which reduces disease activity and damage without hampering normal growth, development and fertility is still an elusive one.

Microbial cultures:

National Center for Microbial Resource (NCMR) NCCS, Pune (Size: 180000 micro-orgnisms):

NCMR was established at **National Centre of Cell Sciences, Pune (an Autonomous Institute of DBT)** by DBT in 2008 for the purpose of preservation of microbial diversity of the country. It is recognized as Designated Repository by Ministry of Environment, Forests & Climate Change under the Biodiversity Act 2002 and International Depositary Authority under the Budapest Treaty. It has world's largest collection of microbes under single roof and offers variety of identification and storage services. It plays a crucial role as custodian of microbial diversity of India. With more than 180000 microorgnisms in its collection NCMR is the largest culture collection in the world and single-handedly lifted India to 3rd place among countries having collection of microorganisms. It has world class expertise in the maintenance, identification and preservation of microbes. It also has well laid out systems and procedures or acquisition of microbial sources from depositors which can be easily translated for obtaining resistant microbes. Considering AMR as the top most national priority, the Department has notified National Centre for Microbial Resource (NCMR) to function as a Bio-repository for resistant microbes/infective agents (Bacteria and Fungi)" and to carry out collection, storage, maintenance, preservation and characterization of these microbes across the country.

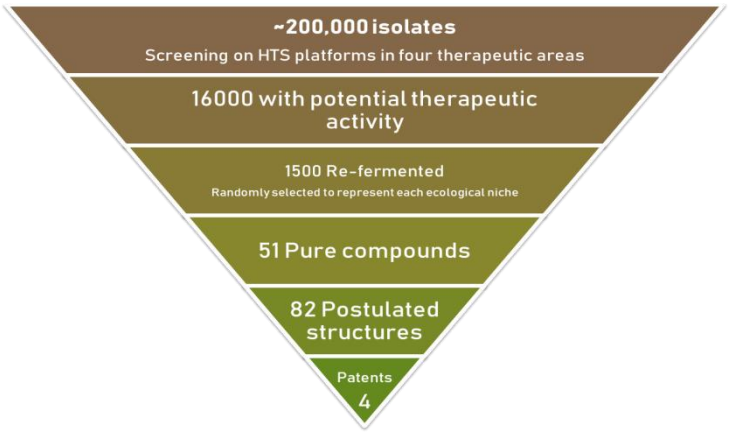


Fig.: Current Status of Cultures obtained from Bio-prospecting Projects

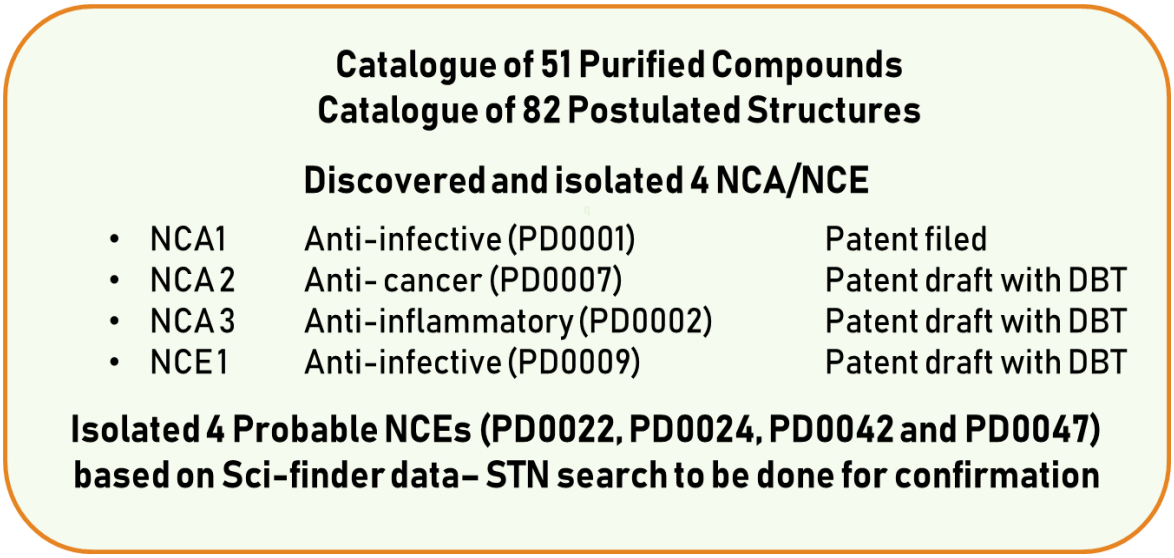


Fig.: Current status of leads obtained from Bio-prospecting Projects

References

1. Loft S, Poulsen HE. Cancer risk and oxidative DNA damage in man. J Mol Med (Berlin, Germany). 1996;74:297–312.
2. Carey DJ, Fetterolf SN, Davis D, Faucett WA, Kirchner HL, Mirshahi U, Murray MF, Smelser DT, Gerhard GS, Ledbetter DH. The Geisinger MyCode community health initiative: an electronic health record-linked biobank for precision medicine research. Genet Med. 2016;18:906–13.
3. Kumaran K, Yajnik P, Lubree H, Joglekar C, Bhat D, Katre P, Joshi S, Ladkat R, Fall C and Yajnik C. The Pune Rural Intervention in Young Adolescents (PRIYA) study: design and methods of a randomised controlled trial. BMC Nutrition (2017) 3:41.
4. Vaught J, Kelly A, and Hewitt R. A Review of International Biobanks and Networks: Success Factors and Key Benchmarks. Biopreservation & Biobanking. Volume 7, Number 3, 2009/2010