

# Golestan Cohort study for Esophageal Cancer

Cohorts and Consortia Conference. Banff Canada 2009

**Reza Malekzadeh M.D. AGAF**

**Professor of Medicine**

**Tehran University of Medical Sciences**

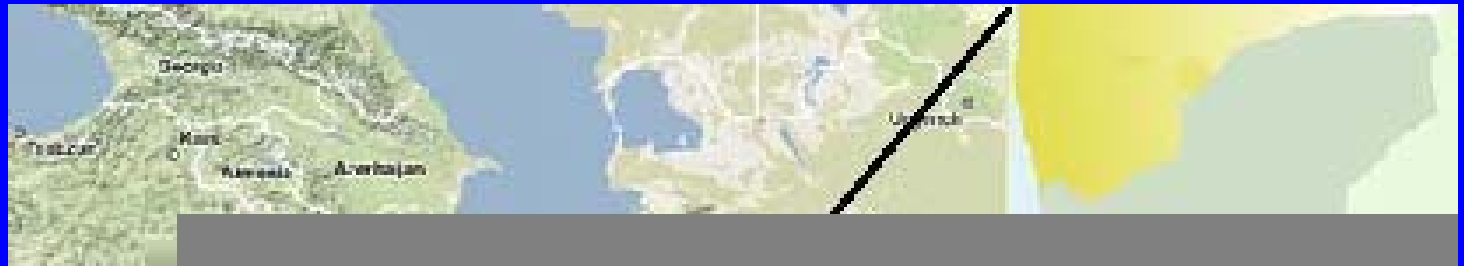
**Tehran Iran**



# North Eastern Iran with highest SCC rates in the world

- Incidence rates of **109** out of **100 000** among **men** and **174** out of **100 000** among **women** (Kmet and Mahboubi, 1972; Mahboubi et al, 1973).
- Sharp changes in the incidence of EC, 300 km to the West of the country .
- A series of studies in the 1970s were prematurely stopped in 1978, and though not conclusive in explaining the increased risk

# Catchment area



# Esophageal Squamous Cell Cancer

## Environmental Risk Factors

---

**Low Risk**

**Tobacco**

**Alcohol**

**Diet**

**Low SES**

**High Risk**

**Hot Liquids**

**Diet**

**Opium**

**PAHs ?**

**Nitrosamines ?**

**Family History**

**Fungal toxins ?**

**All Areas**

**Head & Neck Ca**

**Achalasia**

**Caustic injuries**

**Celiac disease**

**Tylosis**

**Plummer-Vinson**

# Gastroesophageal Malignancy in Northern Iran (GEMINI) Project

2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2014...

**Cancer Registry**

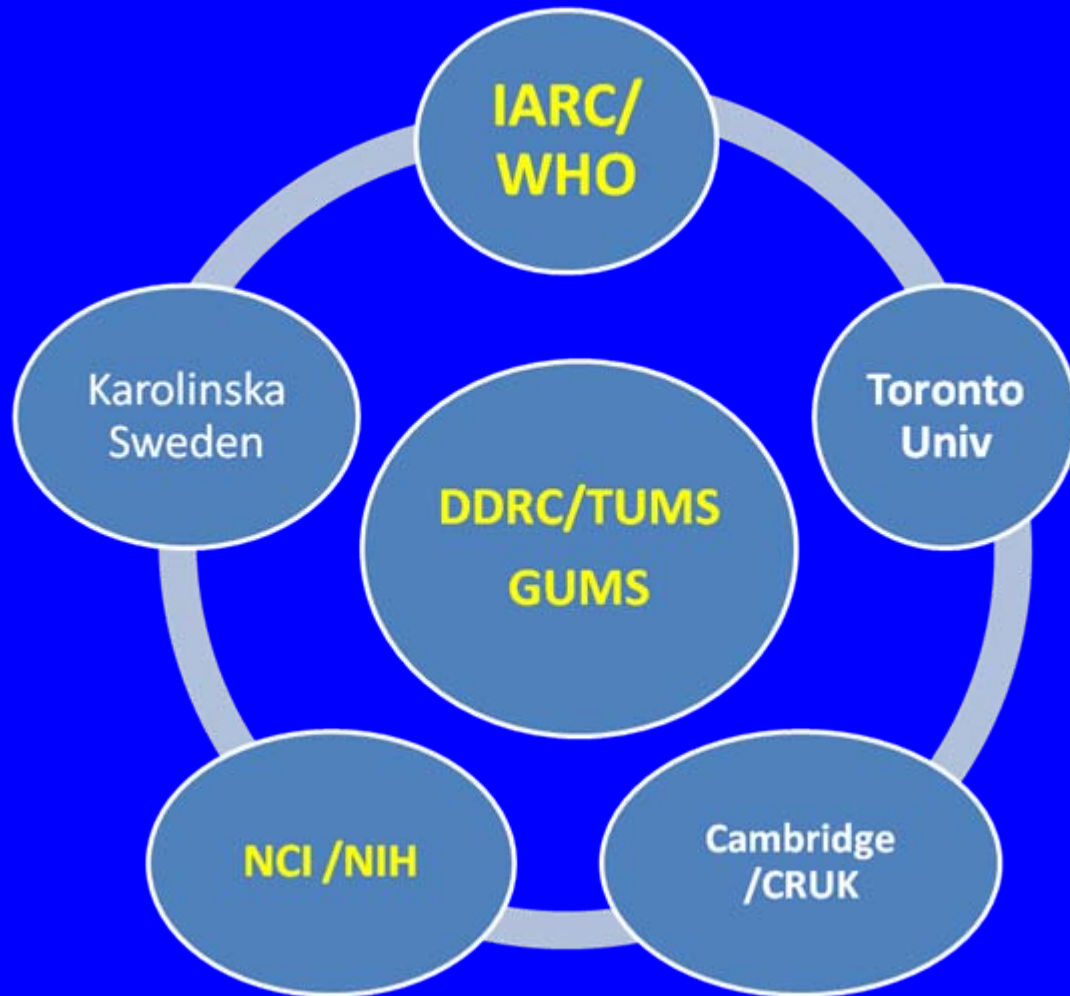
**Case-Control, Ecologic**

**Molecular Pathology/Genetics**

**Cohort**

**Early Detection/Screening  
Prevention**

# International Collaborations





# **Atrak Clinic**

## **Gonbad City Golestan Province**

# Atrak Clinic :Providing Free service to Cohort member with Upper GI symptoms

--	--

# Aims

- To identify risk factors for EC by a assessment of
- Tobacco, opium and alcohol use.
- Tea drinking habits, including tea temperature
- Ethnicity, physical activity
- Socioeconomic status, Occupational history,
- Family history of cancers,
- Gastrointestinal symptoms and signs,
- Anthropometric characteristics ,Oral health.

-

# Evaluation of Role of Nutrition

- A Validated food frequency questionnaire (FFQ) specifically **developed** for this population and validated during the pilot study 18. The FFQ covers 116 food items, including bread and cereals, meat and dairy products, oils, sweets, legumes, vegetables, fruits, and condiments, as well as cooking methods.

# Biospecimen banks

- To establish biospecimen banks for blood, urine, hair, and nail samples to be used in molecular and genetic studies of cross-sectional or nested case-control design.

# Who is in the sample?

## Size of Cohort: 51000 subjects

- The study population is a sample of the Golestan population, aged 40-75 years.
- 10,032 urban participants were enrolled with a participation rate of 70% for women and 50% for men
- 40,013 participants were enrolled from 326 villages, with participation rates of 84% for women and 70% for men.








\*Discussing the study

\*Taking a consent

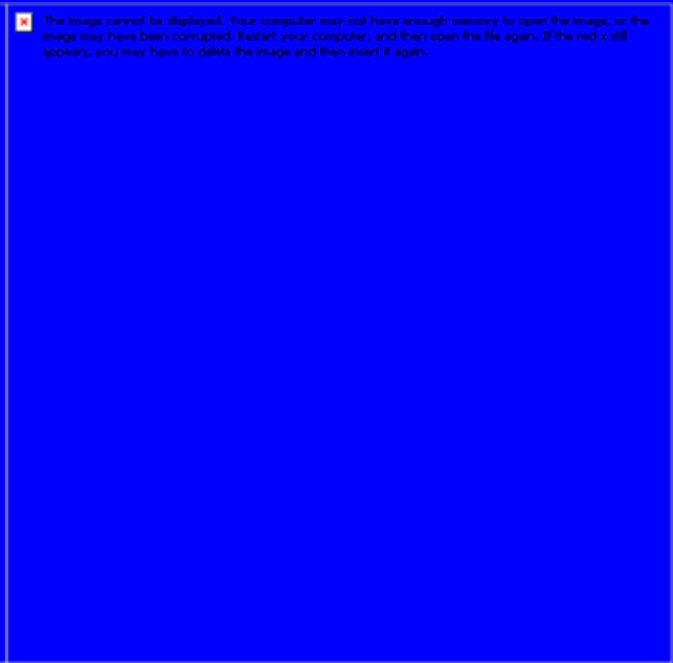
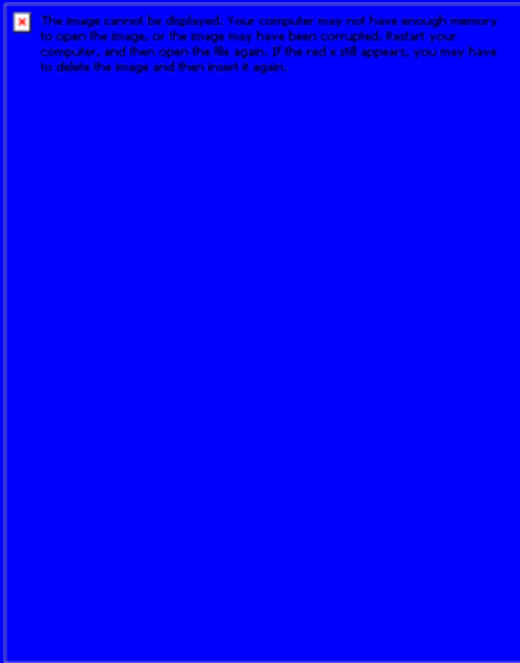
\*Recording participants' general  
information

# GP interview


 The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

A nutritionist is interviewing using FFQ  
( a localized food portion sized album is used)


# Health workers measuring H, W, ..., tea temperature & assessing Socio-economic



# Blood, urine, nail, and hair sampling

 The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

 The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

 The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

 The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.











# Percentage of cohort participants who provided biological samples

- Almost 98% of subjects provided blood
- More than 99% provided, urine, hair, ND nails.
- Tissue Blocks will be available for at least 60 percent of cases with cancer diagnosis.

# Anything new is likely to be learned about the aetiology of oesophageal cancer from this study.

- Tobacco and alcohol do not seem to be very important in this population and this is the reason the study may provide unique information.
- **Dietary imbalances, hot drinks and other exogenous factors** are more important.
- Most of these factors can only be adequately studied prospectively.

# Initial list of hypotheses that can be tested in this cohort

- Tea temperature (from general questionnaire)
- Infections (Helicobacter Pylori, HPV....)
- Low intake of fruit and vegetables (from FFQ)
- A dietary PAH (polycyclic aromatic hydrocarbons) from 1-OH-pyrene in urine.
- Common Genetic cancer susceptibility variants

# **Secondary analyses for which the cohort will have enough power**

- **BMI and overall/CVD mortality;**
- **smoking and major causes of death,**
- **Viral hepatitis(HBV,HCV) liver mortality.**
- **COPD and lung related mortality**
- **Risk factors for CVA**

# Methods of follow up

- **Active telephone contact.**
- **Active follow-up by team .**
- **Passive follow-up .**
- **Follow-up by Auxiliary health workers (Behvarz )**
- **The Golestan cancer registry**
- **Atrak clinic**
- **Verbal Autopsy**

# *Follow up*

- Completing the FU questionnaire by telephone call

# *Follow up*

- Capturing photos of all documents from participants who encounter an outcome

# 372 incident cancers identified to date

- 129, **esophageal** cancer and **gastric cancer** together comprised **35%**
- Colorectal, breast, and lung cancers together comprised only **19.6%** of the total.
- **Liver**                    **5%**
- Leukaemia            5 %
- Pancreas            4%
- Ovarian            3.5%
- Lymphomas.    3%
- Others less than 1 percent of the total.

# Follow-up Report

The total number of Deaths, has reached to 1147 deaths. It means that the mortality rate is 994 in 100000 person years. The distribution of causes of death is as below;

Cause of Death		Number (%)	
Cardiovascular	MI	351	385 (33.57%)
	Other	34	
Cancer		233 (20.31%)	
CVA		180 (15.69%)	
Pulmonary related		67 (5.84%)	
Vehicle Accident		46 (4.01%)	
Renal		24 (2.09%)	
Liver related		24 (2.09%)	
Other		135(11.77%)	
Pending		14 (1.22%)	
Unknown		39 (3.40%)	
Total		1147 (100%)	

# Ascertainment of outcomes

- Pathology reports are obtained for 75 percent of cancer cases.
- A combination of clinical records and verbal autopsy is used for the rest.

# Verbal autopsy

- Verbal autopsy is the only method of case ascertainment for less than 20 percent of cases.
- A study comparing the diagnoses made through verbal autopsy with diagnoses made on basis of medical records is currently underway.

# The follow up success rate

- First year: 99.7%
- Second year: 99.8%,
- Third year: 99.8%
- Fourth year: 99.8%
- Fifth year : 99.8%.

# Unique distributions of cancer diagnosis

- The unique distributions of cancer diagnosis and environmental exposures in this cohort may enhance the opportunity for discovering new risk factors

# A Model for Cancer Research in developing Countries\*

- Most cancer epidemiology studies involve people living in North America and Europe, which represent only a fraction of the global population.

# New Dimensions in Cancer research

- The wide variety of dietary, lifestyle and environmental exposures, as well as the genetic variation among people in developing countries can provide valuable new information on factors that contribute to cancer or that protect against it.

# A Model

- To provide a model for population-based studies in a country in economic and social transition based on collaboration between local health workers, local health authorities, national research centers, and international research institutions

# knowledge of cancer etiology

- These types of studies would not only serve the needs of people in developing countries, but also progress our overall knowledge of cancer aetiology.

# **Power of Study for planned molecular studies**

- **The power for evaluating the main effect of genetic risk factors for cancers will be low during the next five years of follow-up.**
- **Therefore In addition to the cohort study we continue to recruit cancer cases for our CCS in the Atrak clinic - we have already accrued more than 600 cases of oesophageal SCC and 400 of gastric cancer and 50 cases of EAC.**

# CCS cases for Genetic studies

- We hope to recruit over 2,000 UGI cancer cases for genetic studies in next 5 years.
- We will be in a strong position to contribute to future multi-centre, international collaboration in genetic association studies

# **Plans to update the exposure information**

- **Short questions about major exposures for all cohort members every 3 years of follow up.**
- **More detail questioner including FFQ for 10% of all cohort subjects after 5 year of follow up**

End

Slides to answer question and  
comments

# How you made decisions about populations and scientific questions to include/exclude ?

- We have enrolled a population with one the highest rates of esophageal cancer in the world.
- We have enrolled all subjects >40 year of age to be able to have enough cancer at the 5 year and 10 year follow up.

# Resource allocation

- How you decided about resource allocation (multiple blood specimens over time and re-interview versus higher initial numbers and more sparse active follow-up);

# How other disciplines can/did get involved;

- They should be able to fit to the study without interfering with our primary goal.

# How others can be brought in to the research group.

- **CVD** epidemiology group invited : They have finished a pilot phase study of polypill trial for prevention of CVD.
- We have invited our **pulmonologist** to study the COPD in a population of our cohort.
- We have invited **nephrologists** to look for micro-albuminuria in a group of subjects.
- **Hepatologists** already started to establish their cohort of HBV and HCV chronic disease



# Opportunities

- A very good Infrastructure for study of chronic disease of almost all types.
- Best Data source and biological samples for Ph.D students research.
- Possibility of interventional studies for prevention of chronic diseases (CVD,Cancer..)
- Providing the best type of evidence to answer major question about health issues (Should we treat asymptomatic >40 year old cases for HCV infection?)

# Challenges

- How much we can disturb and invite our cohort subjects?
- What are the best and efficient way to use our resources (biobank..)
- What are the priorities for inviting or accepting other groups to work on our cohort?

## **Limitations of the currently available evidence linking diet to oesophageal cancer.**

- Publications from the **EPIC**, NIT cohort in China, the ATBC cohort from Finland, and NIH-AARP cohort provide some data that fruit and vegetable consumption are protective and that red meat consumption is harmful.

# Hypothesis generated from CCS and Pilot phase of GCS.

- Severe deficiency in vitamin intake among women and marked differences in nutrient intake between rural and urban dwellers may contribute to the observed epidemiological pattern of EC in Golestan, with high incidence rates among women and people with low socioeconomic status, and the highest incidence rate among rural women.

# Tea temperature

- Tea temperature was found to have the highest odd ratio in our case control study . We have measured drinking tea temperature in all 50,000 subject and for the first time we have prospective data to retest the tea temperature hypothesis.

# PAH

- Exposure to dietary polycyclic aromatic hydrocarbons may be important.
- We also collected urine from all subjects for possible measurement of PAH metabolites and collected serum and tissue to look for PAH adducts in serum and tissue.

**Degree of dietary heterogeneity  
between study participants**

# Proportion (%) of subjects with daily intakes lower than RDAs or LTIs

	Urban men	Urban women	Rural men	Rural women
<b>Protein</b>				
< RDA (0.75 g/kg body weight)	5.9	52	6.1	38
< LTI (0.45 g/kg body weight)	0	9.1	0	4.2
<b>Vitamin A</b>				
< RDA (700 RE µg for men, 600 RE µg for women)	71	61	76	94
< LTI (300 RE µg for men, 250 RE µg for women)	5.9	3	36	67
<b>Vitamin C</b>				
< RDA (60 mg for both genders)	41	64	97	100
< LTI (32 mg for both genders)	5.9	21	42	73

# Adjusted<sup>1</sup> geometric mean differences (95% CIs) for selected food groups and nutrients after stratification by place of residence and ethnicity

Food groups/ nutrients	Rural Turkmen - Rural Non-Turkmen <sup>2</sup>	<i>P</i> Value	Urban Turkmen - Urban Non- Turkmen <sup>2</sup>	<i>P</i> Value	Urban Turkmen - Rural Turkmen <sup>2</sup>	<i>P</i> Value
Bread (g)	-231 (-456 – -76)	<0.01	-58.2 (-142 – 4.16)	0.07	-136 (-189 – -91.5)	<0.01
Legumes (g)	-0.29 (-25 – 5.2)	0.94	-14.5 (-33.4 – -3.32)	0.01	6.57 (-0.17 – 9.96)	0.05
Vegetables (g)	7.28 (-41.7 – 31.5)	0.69	-49.9 (-87.2 – -20.6)	<0.01	33.7 (13.8 – 48.4)	<0.01
Fruit (g)	36.2 (35.8 – 36.3)	<0.01	30.0 (-54.1 – 82.8)	0.41	147 (95.2 – 168)	<0.01
Meat (g)	17.6 (7.99 – 20.9)	0.01	6.71 (-17.1 – 21.6)	0.51	25.1 (11.98 – 33.43)	<0.01
Carbohydrate (g)	-54.5 (-98.4 – -14.9)	0.01	-11.5 (-26.9 – 3.17)	0.12	-23.7 (-39.9 – -8.24)	<0.01
Fat [nutrient] (g)	25.6 (15.5 – 32.8)	<0.01	6.19 (-1.12 – 12.6)	0.09	9.75 (1.78 – 16.6)	0.02
SFA (g)	6.76 (4.83 – 7.97)	<0.01	2.18 (-0.08 – 4.09)	0.06	5.01 (2.80 – 6.82)	<0.01
MUFA (g)	11.1 (7.28 – 13.8)	<0.01	1.83 (-1.47 – 4.65)	0.25	2.19 (-1.31 – 5.15)	0.20
PUFA (g)	7.44 (4.09 – 9.72)	<0.01	1.63 (-1.43 – 4.12)	0.27	0.48 (-2.25 – 2.79)	0.71
Cholesterol (mg)	49.5 (18.5 – 60.2)	0.01	12.8 (-33.9 – 47.7)	0.54	88.8 (51.9 – 112)	<0.01