

## Original Article



# Building a Cancer Biobank in a Low-Resource Setting in Northern Iran: the Golestan Cancer Biobank

Fatemeh Ghasemi-Kebria, PhD candidate<sup>1\*</sup>; Nastaran Jafari-Delouie, MSc<sup>1\*</sup>; Taghi Amiriani, MD<sup>1</sup>; Alireza Norouzi, MD<sup>1</sup>; Behnoush Abedi-Ardekani, MD<sup>2\*</sup>; Dariush Nasrollahzadeh, PhD<sup>3,2\*</sup>; Mohammad Ashaari, MD<sup>4</sup>; Sima Besharat, PhD<sup>1</sup>; Mohammad Naeimi-Tabiei, MD<sup>5</sup>; Isen Gharanjic, MSc<sup>1</sup>; Zahra Babapalangi, MSc<sup>1</sup>; Hossein Poustchi, PhD<sup>3</sup>; Shahryar Semnani, MD<sup>6,1</sup>; Abdolreza Fazel, MD<sup>5</sup>; Zisis Kozlakidis, PhD<sup>7\*</sup>; Elisabete Weiderpass, PhD<sup>8\*</sup>; Gholamreza Roshandel, PhD<sup>1\*</sup>

<sup>1</sup>Golestan Research Center of Gastroenterology and Hepatology, Golestan University of Medical Sciences, Gorgan, Iran

<sup>2</sup>International Agency for Research on Cancer (IARC), World Health Organization (WHO), Lyon, France

<sup>3</sup>Digestive Oncology Research center, Digestive Disease Research Institute, Tehran University of Medical Sciences, Tehran, Iran

<sup>4</sup>Department of Pathology, Golestan University of Medical Sciences, Gorgan, Iran

<sup>5</sup>Cancer Research Center, Golestan University of Medical Sciences, Gorgan, Iran

<sup>6</sup>Omid Cancer Research Center, Golestan University of Medical Sciences, Gorgan, Iran

<sup>7</sup>Laboratory Service and Biobank Group, International Agency for Research on Cancer, World Health Organization (WHO), Lyon, France

<sup>8</sup>Office of the Director, International Agency for Research on Cancer (IARC), World Health Organization (WHO), Lyon, France.

## Abstract

**Background:** We aim to present the development and the initial results of the Golestan Cancer Biobank (GoCB), in a low resource setting in northern Iran.

**Methods:** The GoCB protocol and its standard operation procedures (SOP) were developed according to internationally accepted standards and protocols with some modifications considering the limited resources in our setting. The main biological samples collected by the GoCB include blood sample, urine sample, fresh endoscopy tissue sample, fresh surgical tissue sample and formalin fixed paraffin embedded (FFPE) tissue sample. The GoCB collects patients' demographic data, tumor characteristics as well as data on risk factors. We developed a specific GoCB software for management of patient data and biological sample information. The GoCB dataset is annually linked with the Golestan cancer registry dataset to add complementary data (e.g., survival data).

**Results:** The GoCB started collection of data and biological samples in December 2016. By November 2020, a total number of 1217 cancer patients participated in the GoCB. The majority of the GoCB participants (n=942, 77%) were those with gastrointestinal and breast cancers. Data on risk factors were successfully collected in 684 (56.2%) of the participants. Overall, 3563 samples were collected from the GoCB participants and 730 samples were used in 7 national and international research projects.

**Conclusion:** We considered specific strategies to overcome major limitations, especially budget shortage, in the development and maintenance of a cancer-specific biological repositories in our setting. The GoCB may be considered as a model for the development of biobank in low- and middle-income countries (LMICs).

**Keywords:** Biobank, Cancer, Iran, Low- and middle-income countries

**Cite this article as:** Ghasemi-Kebria F, Jafari-Delouie N, Amiriani T, Norouzi A, Abedi-Ardekani B, Nasrollahzadeh D, et al. Building a cancer biobank in a low-resource setting in northern Iran: the Golestan Cancer Biobank. Arch Iran Med. 2021;24(7):526-533. doi: 10.34172/aim.2021.75

Received: January 15, 2021, Accepted: March 3, 2021, ePublished: July 1, 2021

## Introduction

Despite the development of oncology and cancer genomics research, cancer is the second leading cause of death worldwide. The latest global cancer data from the International Agency for Research on Cancer (IARC) suggested that cancer burden increased to 19.3 million new cases and 10.0 million cancer deaths in 2018.<sup>1</sup>

The growing cancer incidence and mortality, especially in low- and middle-income countries (LMICs), are due to several factors, including population growth and aging

as well as changes in the prevalence and distribution of the main risk factors.<sup>1,2</sup> Conducting cancer research to identify risk factors is one of the most important steps in designing effective cancer control programs. Identifying risk factors requires the analysis of large numbers of biological samples in order to be able to effectively differentiate causal biological relations. Harmonized and standardized collections of biological materials and associated data for the longer term, and representative of populations, may ensure achieving this goal while

\*Corresponding Author: Gholamreza Roshandel, PhD; Golestan Research Center of Gastroenterology and Hepatology, Sayyad Shirazi hospital, Gorgan, Iran. Tel: +98 911 375 7327; Email: roshandel.g@gmail.com

Abdolreza Fazel, MD, Cancer Research center, Sayyad Shirazi hospital, Gorgan, Iran. Tel: +98 17 32424540; Email: fazelabdolreza@gmail.com

#Contributed equally as first authors

‡Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

promoting cancer research.<sup>3,4</sup> In the current era of availability of sophisticated analytical technologies, the high quality of collected biological material plays a key role in ensuring experimental reproducibility and eventually meaningful interpretation of cancer research. Therefore, the standardized collection of high-quality, research-ready samples and data will help to expand medical research on cancer and will consequently inform cancer control planning.<sup>5</sup>

Biobanking is a resource intensive undertaking, and thus might be overlooked in systems with many competing priorities, especially within LMIC contexts, where it may seem too difficult or even impossible to set up such research infrastructure.<sup>6,7</sup> However, the appropriate adaptation of the standard operation protocols and guidelines according to these limitations can provide a workable technical solution. At the same time, the consideration of the potential benefit(s) of the supported research and the downstream impacts of this research on community health may justify biobanking activities and allow them to be considered as one of the priorities in health-related research.<sup>8</sup>

The Golestan province in Northern Iran has been known as a high-risk area for upper gastrointestinal (GI) cancers since the 1970s.<sup>9,10</sup> The Golestan Population-Based Cancer Registry (GPCR), a voting member of the international association of cancer registries (IACR) since 2007, covers the total population of Golestan and provides high-quality population-based cancer data to researchers as well as to health policy makers at local and national levels since 2004.<sup>11</sup> As reported in 2014,<sup>12</sup> the age-standardized incidence rate (ASR) (per 100 000 person-year) of esophageal cancer in Golestan males (15.81) and females (11.71) was considerably higher than the national rates of males (6.47) and females (5.13). The ASRs of stomach cancer were also higher in the Golestan male (24.79) and female (10.95) populations compared to those reported for Iranian males (21.24) and females (9.44).<sup>12</sup> In addition, recent reports from the GPCR showed increasing trends in the incidence rate of other cancer such as colorectal,<sup>13</sup> breast<sup>14</sup> and lung cancers. Therefore, because of the high rate of gastrointestinal (GI) cancers and the increasing incidence trend of other cancers in the Golestan province,<sup>15</sup> promotion of cancer research activities has been taken into account by the medical research communities and policymakers in this area. For this purpose, the Golestan Cancer Biobank (GoCB) was established in 2016 to provide biological samples and associated data collected from cancer patients by applying international standards and supporting high-quality cancer research. This article will present the process of designing and creating this biobank infrastructure (within an LMIC context), as well as the initial results of the GoCB, as the first and only cancer-specific biobank in the Golestan province, a high-risk area in northern Iran.

## Materials and Methods

### GoCB Protocol

The creation of GoCB was conducted as a 5-year

research project. As the first step, the GoCB protocol and corresponding standard operation procedures (SOP) were developed considering internationally accepted standards and protocols<sup>16,17</sup> and for different operational aspects of the GoCB, such as the setting up of the GoCB lab, the GoCB field centers, types of samples, sample collection and sample processing methods, sample preservation and the access policy for samples.

### GoCB Organization

The GoCB organization chart includes a steering committee, a scientific committee and the GoCB secretariat. The steering committee consists of the principal investigators (PIs) of the GoCB project. The scientific committee consists of clinicians and researchers specializing in various fields related to cancer and biorepositories, including cancer surgery, oncology, pathology, genetics, biochemistry, cellular and molecular biology, and epidemiology. The main activity of the GoCB scientific committee is to supervise all scientific aspects of the GoCB, especially the development and revisions of the main protocol and the specified SOPs. The GoCB secretariat consists of dedicated trained staff, located in the Golestan Research Center of Gastroenterology and Hepatology (GRCGH) affiliated with the Golestan University of Medical Sciences (GOUMS), Gorgan, Iran.

### Setting up the GoCB Laboratory

The GoCB lab, at the GRCGH, is located within Sayyad Shirazi hospital, Gorgan city, northern Iran. It is a referral tertiary public hospital for cancer patients of the Golestan province and other cities of neighboring provinces, including Semnan and North Khorasan provinces. The GoCB lab was equipped from the outset with standard laboratory cabinets (re-purposed), centrifuges, vertical laminar flow hood as well as equipment to process the tissues into formalin fixed paraffin embedded (FFPE) tissue blocks. The GoCB lab was also equipped with -80°C freezers for storage of fresh tissue samples as well as blood and urine samples. An in-house GoCB-specific, FFPE box and cabinet were created for the storage of FFPE samples (Figure 1).

### GoCB Field Centers

The GoCB also receives biological samples from cancer patients from its field centers (sample collection sites) which are mainly located in public and private hospitals. Additional selected outpatient centers (e.g., centers providing endoscopy, colonoscopy, bronchoscopy and colposcopy services) also collaborate with the GoCB as potential field centers. In each GoCB field center, an experienced healthcare worker (nurse, laboratory assistant, midwife, surgical operation assistant and endoscopy/colonoscopy/bronchoscopy assistant) is selected and trained to be assigned as GoCB field staff. These GoCB field staff are full-time personnel of their field centers and are compensated for their efforts proportionately, depending on the number of participants they recruited. As part of continuous training program,



**Figure 1.** Golestan Cancer Biobank (GoCB) Specific Box (A) and Cabinet (B) for Preservation of Formalin-Fixed Paraffin-Embedded (FFPE) Tissue Samples.

GoCB annual meetings are arranged during which all GoCB staff are invited for reviewing and open discussion on the protocols, as well as to appreciate their contribution to the GoCB.

### *Types of Biological Samples*

The main biological samples collected by the GoCB include blood sample, urine sample, fresh mucosal biopsies by endoscopy, fresh surgically resected tissue samples and FFPE tissue samples. GoCB systematically collects tumor/non-tumor pairs from all cancer patients, if routine diagnostic sample collection is not hampered.

### *Collection of Data*

After obtaining informed consent, data collection is performed using two GoCB-specific data collection forms. The GoCB data collection form 1 (general questionnaire) consists of patients' demographic data, clinical data and tumor characteristics (site of tumor, morphology, behavior and tumor grade); the latter are collected based on the pathology report. The third edition of the International Classification of Diseases for Oncology (ICD-O-3) is used for coding tumor characteristics, including topography, morphology, behavior and grade.<sup>18</sup>

The GoCB data collection form 2 (risk factor questionnaire) was developed using the Golestan Cohort Study questionnaire.<sup>19</sup> It covers a broad range of generic risk factors applicable to all cancer types as well as specific ones based on the epidemiological studies on common cancer types in our population, including

anthropometric measures, level of education, occupation, smoking behavior, alcohol and opium consumption, diet, oral hygiene habits and medical history, cancer history in family, exposure to X-ray radiation, exposure to pesticides, physical activity and data on reproductive behaviors (only for women).

### *Collection of Biological Samples*

The GoCB guidelines for the collection of biological samples include detailed information on each type of biological sample collected, including temperature, time and place of preservation of sample in the filed center, condition of transporting the sample to the GoCB lab, as well as name and telephone number of contact persons in the GoCB laboratory. Additional information includes ischemia time (for surgical resected samples), date and time of sample shipment to the GoCB and cool box temperature during sample shipments.

Following the patient's informed consent, blood samples (9 mL) are collected using ethylenediaminetetraacetic acid tubes. Blood samples are shipped to the GoCB laboratory within 2 hours after collection using a cool box and ice pack to ensure appropriate temperature conditions during shipment. Mid-stream urine samples are collected using specific sterile containers and transported to GoCB laboratory in the same package and transport facility as blood.

Fresh surgical tissue samples (size of 0.5 cm \* 1 cm \* 1 cm) and fresh endoscopy tissue samples are collected using containers with RNA-later and stored at room temperature until shipment to the GoCB laboratory. Formalin fixed surgical tissue samples are collected using prefilled containers with 10% buffered formalin and stored at room temperature until shipment to the GoCB laboratory.

### *Quality of Collected Samples and Data*

Upon receiving the biological samples and questionnaires, the quality of data and biological samples are checked by GoCB secretariat staff to ensure adherence to the GoCB guidelines and protocols. An in-house developed quality control checklist allows to check for the completeness and accuracy of data in a consistent manner. If there is any deviation observed, e.g., missing data, errors in filling out the questionnaire, unclear data, inadequate amount of specimen (volume or size) and errors in labeling, the GoCB secretariat staff provide feedback to the field staff and follow up the issue until resolved.

### *Processing and Preservation of Biological Samples*

GoCB uses high-quality cryotubes (of 0.5, 1.4 and 3 mL) according to the type of biological sample. The tubes are sterile and free from DNase, RNase, DNA, and endotoxins. Each tube has a unique ID, as a laser etched 2-D barcode on the base. A 2-D barcode reader is used to ensure entering the correct sample code into the GoCB software.

### *Processing and Preservation of Blood Sample*

Blood samples are separated and stored in 1 mL aliquots,



as different parts including whole blood (red cap), plasma (white cap), Buffy coat (blue cap) and red blood cell (RBC) (gray cap).<sup>16</sup> Urine samples are also stored in 1 mL aliquots. The samples information is registered in the GoCB software through the sample barcode, as well as detailed information on the sample location (freezer, rack, box, place within the box). After registering in the GoCB software, the samples are transferred and stored in the -80°C freezer.

### Processing and Preservation of Fresh Tissue Samples

After receiving fresh endoscopy and surgical tissue samples in the GoCB laboratory, tissue containers with RNA-later are placed overnight in refrigerator (4°C). On the next day, the tissue samples are removed from the container and RNA-later and transferred into 0.5 mL cryotube (endoscopy sample) or 3 mL cryotube (surgical resected sample) with external thread screw caps. After registering sample information in the GoCB software, the samples are transferred and stored in the -80°C freezer.

### Processing and Preservation of FFPE Samples

After receiving formalin fixed tissue samples, the tissue specimens are processed and embedded in paraffin in accordance with routine histological techniques for the preparation of FFPE blocks.<sup>20</sup> The FFPE samples are located in the GoCB-specific FFPE box and stored in our specific FFPE cabinet at room temperature.

### GoCB Software and Dataset

An in-house software was developed for management of patient data and biological samples information (Figure 2). The GoCB software consists of two main interfaces: the

sample registration interface and the patient registration interface. The two interfaces are linked using a unique GoCB registration number.

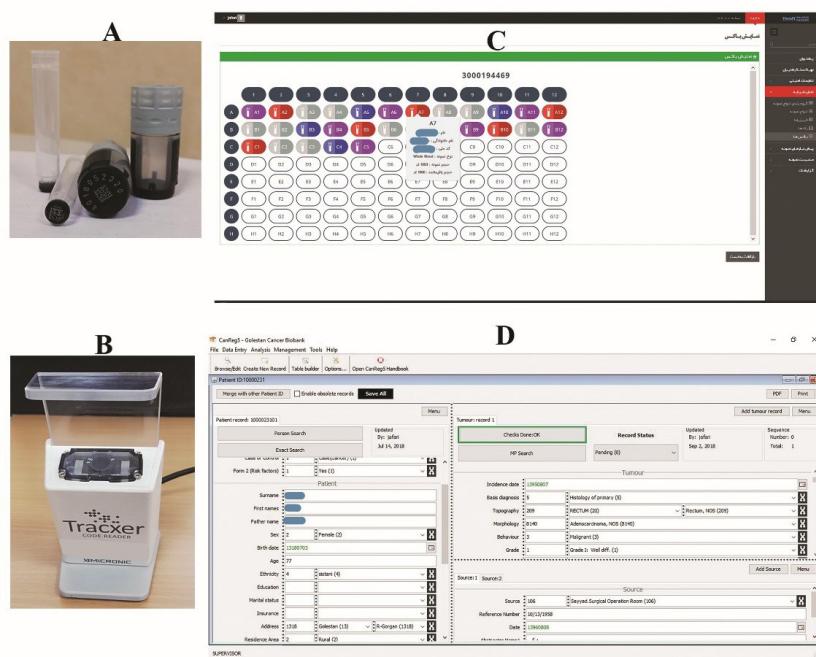
The sample registration interface is a laboratory information management system, called Bio-Tracker, version 3.3.0, locally developed by “Batab Tajhiz Parsin” Company, Iran. It contains detailed information on the biological samples.

The patient registration interface consists of detailed information on the patients including demographic data, clinical data as well as risk factor data, as collected by the standardized questionnaires. The patient registration interface of the GoCB software was developed using the CanReg5, an open-source tool developed by the IARC to input, store, check and analyze cancer registry data.<sup>21</sup> The CanReg5 tables, fields and variables were modified and aligned to the GoCB questionnaires and prepared a GoCB specific dataset in the CanReg5.

Importantly, using the CanReg5 facilities (providing unique codes for patient, tumor and source) and considering patients’ national identification numbers, all biological samples of a patient collected from different sources (field centers) or at different times and even from different tumor sites (in patients with multiple primary tumors) are linked to the patient’s file using a unique GoCB registration number.

### Linkage between the GoCB Dataset with Golestan Cancer Registry Dataset

The GPCR is a high-quality cancer registry which covers the entire Golestan population and collects population-based cancer data since 2004.<sup>11,15</sup> The GoCB dataset is annually linked with the GPCR dataset to add additional patients’ data (e.g., survival data, etc.) to the GoCB dataset.



**Figure 2.** Software and Equipment for Registration and Management of Patients’ Data and Biological Sample Information in Golestan Cancer Biobank (GoCB). A, Cryotubes with 2-D barcodes; B, Barcode reader; C, Biological sample registration interface; D, Patient registration interface.

### Use of the GoCB Sample and Data

The GoCB samples and data may be provided to researchers according to a research proposal. The proposals are submitted by email to the GoCB secretariat (gocb@goums.ac.ir). Only pseudo-anonymized samples and data will be available, after approval of the proposal in the GoCB scientific committee. The requested samples and data will be provided to the PI of the research project by signing a material transfer agreement.

At the time of delivering selected samples to research PIs, the GoCB secretariat may also provide the list of physicians involved in the collection of those samples (considering the number of samples), while both physician and research PIs are blinded regarding personal information of the research participants. This will help the PIs to access the physicians if they need further information and help regarding the clinical aspects of the disease and specimens and also regarding the interpretation and clinical application of the research findings. In addition, the physicians may receive (via the GoCB) feedbacks from the research team regarding the samples and research findings, and it may also give them a chance to be actively involved and collaborate in research projects.

### Ethical Considerations

Participants must complete an informed consent form before starting sample and data collection. All information about the project including its benefits and risks as well as indications of using the samples, especially possible use of the samples in international research project are explained to the patient. All patients are well informed about their right to refuse or stop their participation at any stage of the project. The GoCB secretariat and field staff as well as clinicians are trained to ensure that participation in the GoCB does not interfere with appropriate patient diagnosis or care. The GoCB uses anonymized codes to ensure the protection of individuals' privacy. The GoCB secretariat and staff have limited access to the dataset according to their role and only a representative of the GoCB steering committee (GR) is authorized to have full access to the GoCB dataset. Neither physicians nor patients are reimbursed for participation in the GoCB.

### Results

The GoCB started collecting data and biological samples in December 2016. Between December 2016 and November 2020, a total of 1256 cancer patients were invited to participate in the project. Thirty-nine of these patients (3.1%) did not consent for sample collection and were therefore excluded. Finally, 1217 cancer patients participated in the GoCB including 484 (39.8%) males with a mean (SD) age of 62.5 (13.3) years and 733 (60.2%) females with a mean (SD) age of 53.2 (14.4) years. Of the total GoCB participants, 715 (58.8%) were residents of urban areas and 502 (41.2%) were from rural settings. [Table 1](#) shows the number and proportions of the GoCB samples by tumor site. The majority of the GoCB participants (n = 942, 77%) were those with gastrointestinal (esophagus, stomach, colorectal) and breast cancers. The

GoCB participants were recruited from 18 different field centers through the contribution of 29 field staff and 23 physicians of 5 different specialties ([Table 2](#)). Complete data sets were successfully collected from 793 (65.2%) of the total 1217 GoCB participants.

Overall, 3563 samples were collected from the study participants between December 2016 and November 2020, as shown on [Table 3](#) by sample type.

By November 2020, a total of 730 GoCB samples were used in 7 national and international research projects, for example the Mutographs project funded by the Cancer Research UK (CRUK) grand challenge program.<sup>22</sup> According to its preliminary findings, two GoCB tissue samples went through standard pipeline for whole genome sequencing of tumor. The quality control steps for data and sample quality suggested that percentage of tumoral cells through using allele-specific copy number analysis was 48% and 60% which is at the level of acceptance for downstream analysis (the average tumor purity in this multi-center studies is around 42%). The findings also

**Table 1.** Number and Percent of Participants of Golestan Cancer Biobank (GoCB) by Tumor Site, 2016–2020

Tumor Site	Number	Percent
Tongue	2	0.16
Parotid gland	1	0.08
Nasopharynx	1	0.08
Esophagus	116	9.53
Stomach	206	16.93
Small intestine	14	1.15
Colon	193	15.86
Rectum	175	14.38
Liver	5	0.41
Gall bladder	19	1.56
Pancreas	16	1.31
Lung	18	1.48
Thymus	1	0.08
Bone	1	0.08
Leukemia	7	0.58
Skin	5	0.41
Soft tissue	6	0.49
Breast	252	20.71
Uterine cervix	5	0.41
Uterus	19	1.56
Ovary	50	4.11
Prostate	4	0.33
Kidney	1	0.08
Bladder	2	0.16
Brain	1	0.08
Thyroid	43	3.53
Lymphoma	5	0.41
Unknown primary site	49	4.03
<b>All sites</b>	<b>1217</b>	<b>100.00</b>

**Table 2.** Number of Participants of Golestan Cancer Biobank (GoCB) by GoCB Field Centers and Physician, 2016–2020

Type of center	Number of Centers/Physicians	Number of Participants	Percent of Participants
Hospital Operation room	3	245	15.01
Hospital inpatient wards	7	512	31.37
ERCP center	1	30	1.84
Endoscopy center	5	481	29.47
Outpatient oncology center	1	26	1.59
Pathology lab	1	338	20.71
Total	18	1632	100
<b>Physicians specialty</b>			
Surgeon	6	549	32.65
Pathologist	5	613	36.45
Gastroenterologist	7	436	25.91
Oncologist	3	75	4.45
Pulmonologist	2	9	0.54
Total	23	1681	100.00

ERCP, Endoscopic retrograde cholangiopancreatography.

**Table 3.** Number of Biological Samples Collected by Golestan Cancer Biobank (GoCB) by Sample Type, 2016–2020

Sample Type	Number	Percent
Blood	1095	30.72
Urine	826	23.19
Endoscopic fresh mucosal Biopsy	424	11.89
Surgically resected fresh tissue	541	15.19
FFPE tissue	678	19.02
Total	3563	100

FFPE, formalin-fixed paraffin-embedded.

suggested that more than 97% of reads were mappable to human genome without report of contamination.

## Discussion

We presented detailed information on the design and operation of the GoCB, a cancer biobank in a restricted resource setting in northern Iran. As in other LMICs, the limitations and challenges were addressed through identifying the best strategies available, based on specific conditions and available resources.

During the design phase, GoCB protocols were developed through close collaboration with national and international institutions and experts and the use of available, validated international standard protocols and guidelines.<sup>16,17</sup> The GoCB has been accepted as a member of the biobank and cohort building network (BCNet) of the IARC as of August 2019,<sup>23</sup> and the International society for biological and environmental repositories (ISBER) in

January 2020.

The GoCB secretariat and laboratory were physically located within a public hospital, as a cost mitigation strategy. By using available facilities in the hospital (building, power supply, air circulation, sanitation, etc.), the GoCB could substantially minimize the setting up and operational costs. In addition, using in-house made equipment (e.g., GoCB-specific FFPE box and cabinet) and development of in-house GoCB software, especially using the open source CanReg5 tool, played a critical role in preserving the limited available resources within our existing settings.

According to the GoCB plan for the use of samples and data, the name of physicians involved in sample collection were provided to the research team(s) and this could establish a link between the physician and the research team, resulting in possible closer collaborations between all parties involved. Such potential collaborations could help research teams to improve the quality of their research and on the other hand, could motivate the physicians to become involved even more actively in research projects through the GoCB. In order to maximize the use of funds in each GoCB field center, the center personnel operating as GoCB field staff collaborate with the GoCB on a part-time basis, while maintaining their main jobs in the field center.

In order to maintain the stakeholder engagement for the longer term, all GoCB collaborators, including physicians and field staff were provided official certificates according to their contributions during the different phases of the GoCB project. As part of the quality maintenance activities, the GoCB also considered continuous professional training and education programs, as well as annual appreciation ceremonies for the GoCB collaborators. In addition, providing regular feedback on the quality of samples and data to the GoCB field centers was another strategy to ensure maintenance of the GoCB strong communication channels.

GI and breast cancers were the most common malignancies recruited in the GoCB during its first four years. This was not surprising, because they are the most common cancers in the Golestan population.<sup>13-15,24,25</sup> According to the latest GPCR report (unpublished data), cancers of the breast (female), stomach, colorectum, lung and esophagus were the most common malignancies in the Golestan population in 2017 ([Table S1, Supplementary file 1](#)).

Access to high-quality population-based cancer registry data is a major strength of the GoCB. By linking the GoCB dataset with the GPCR dataset, we could add complementary data (e.g. vital status and survival data) to the GoCB dataset and this could result in improving the GoCB dataset and could consequently improve the quality of researches on GoCB samples. Collaboration of GoCB with international research projects and networks<sup>22,26</sup> is another strength of the GoCB. These collaborations may result in accessing shared resources (e.g., expertise, protocols) and joining further collaborative works and these will finally improve the quality and promote the

GoCB project.

As GoCB was designed and funded as a research project, lack of continuous financial support in our setting is a major limitation and its maintenance may be affected once the project ends. Therefore, although research grants were available and accessible resources in development of the GoCB, stable and continuous supports are necessary for successful maintenance. This point should be taken into account in all low-resource settings including LMICs.

In conclusion, we considered specific strategies to ensure the development and maintenance of the GoCB as a cancer-specific biological repository regarding limited resources in our high-risk population. Some of the most important strategies included considering research grant as available financial resource, using available infrastructures in public sector, considering re-purposed equipment, developing specific in-house equipment and software, considering specific plans to promote clinicians' contribution to collection of biological samples, assigning selected and trained personnel of healthcare centers as part-time field staff of the GoCB and considering justified incentives for GoCB collaborators. These strategies may be applicable in other similar low-resource settings in Iran as well as in other LMICs. Therefore, the GoCB may be considered as a model for designing and implementation of cancer biobanks in the LMICs. As a research infrastructure, it may play an important role in development of basic cancer research, especially in the field of cancer biology in our high-risk area.

#### Authors' Contribution

FG, NJ: Collaborated in data collection; wrote the manuscript; TA, AN, MA, MN, IG, ZB: Collaborated in data collection; edited and critically reviewed manuscript; BA, DN, SB, HP: collaborated in designing the SOPs; edited and critically reviewed manuscript; EW, ZK, SS: collaborated in quality control; edited and critically reviewed manuscript; AF: initiated, conceptualized and designed the study; edited and critically reviewed manuscript; GR: initiated, conceptualized and designed the study; designed the SOPs; wrote the manuscript; All authors read and approved the final manuscript.

#### Conflict of Interest Disclosures

The authors declare that they have no conflict of interest

#### Ethical Statement

The GoCB project was approved by the research council and ethics committee of the Golestan University of Medical Sciences (ethics code: IR.GOUMS.REC.1394.353).

#### Supplementary Materials

Supplementary file 1 contains Table S1.

#### Acknowledgements

The authors would like to thank Professor Reza Malekzadeh for his great support and supervision during all stages of the GoCB establishment including the design, implementation and maintenance. We also appreciate great contributions of Dr Mehdi Alimadadi, Dr Abdollah Pooshani, Miss Nesa Shokouhifar as well as all clinicians and healthcare workers (the GoCB field staff) in the GoCB field centers for their help and

support in collection of data and biological samples. This project was supported by the Golestan University of Medical Sciences (grant number= 931128254) and (grant number = 700-D-5548).

#### References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209-49. doi: 10.3322/caac.21660.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394-424. doi: 10.3322/caac.21492.
3. Patil S, Majumdar B, Awan KH, Sarode GS, Sarode SC, Gadail AR, et al. Cancer oriented biobanks: a comprehensive review. *Oncol Rev.* 2018;12(1):357. doi: 10.4081/oncol.2018.357.
4. Garraway LA, Lander ES. Lessons from the cancer genome. *Cell.* 2013;153(1):17-37. doi: 10.1016/j.cell.2013.03.002.
5. Lee S, Jung PE, Lee Y. Publicly-funded biobanks and networks in East Asia. *Springerplus.* 2016;5(1):1080. doi: 10.1186/s40064-016-2723-2.
6. Mendy M, Caboux E, Sylla BS, Dillner J, Chinquee J, Wild C. Infrastructure and facilities for human biobanking in low- and middle-income countries: a situation analysis. *Pathobiology.* 2014;81(5-6):252-60. doi: 10.1159/000362093.
7. Mohammadi M. Biobanking in the developing world; maximum specimens, minimum infrastructure. *Basic Clin Cancer Res.* 2017;9(4):1-3.
8. Rotimi C, Abayomi A, Abimiku A, Adabayeri VM, Adebamowo C, Adebisi E, et al. Research capacity. Enabling the genomic revolution in Africa. *Science.* 2014;344(6190):1346-8. doi: 10.1126/science.1251546.
9. Kmet J, Mahboubi E. Esophageal cancer in the Caspian littoral of Iran: initial studies. *Science.* 1972;175(4024):846-53. doi: 10.1126/science.175.4024.846.
10. Mahboubi E, Kmet J, Cook PJ, Day NE, Ghadirian P, Salmasizadeh S. Oesophageal cancer studies in the Caspian littoral of Iran: the Caspian cancer registry. *Br J Cancer.* 1973;28(3):197-214. doi: 10.1038/bjc.1973.138.
11. Roshandel G, Sadjadi A, Aarabi M, Keshtkar A, Sedaghat SM, Nouraie SM, et al. Cancer incidence in Golestan province: report of an ongoing population-based cancer registry in Iran between 2004 and 2008. *Arch Iran Med.* 2012;15(4):196-200.
12. Roshandel G, Ghanbari-Motlagh A, Partovipour E, Salavati F, Hasanpour-Heidari S, Mohammadi G, et al. Cancer incidence in Iran in 2014: results of the Iranian National Population-based Cancer Registry. *Cancer Epidemiol.* 2019;61:50-8. doi: 10.1016/j.canep.2019.05.009.
13. Hasanpour-Heidari S, Fazel A, Semnani S, Khandoozi SR, Amiriani T, Sedaghat S, et al. Temporal and geographical variations in colorectal cancer incidence in Northern Iran 2004-2013. *Cancer Epidemiol.* 2019;59:143-7. doi: 10.1016/j.canep.2019.02.003.
14. Fazel A, Hasanpour-Heidari S, Salamat F, Rajaie S, Kazeminezhad V, Naeimi-Tabiei M, et al. Marked increase in breast cancer incidence in young women: a 10-year study from Northern Iran, 2004-2013. *Cancer Epidemiol.*



- 2019;62:101573. doi: 10.1016/j.canep.2019.101573.
15. Roshandel G, Semnani S, Fazel A, Honarvar M, Taziki M, Sedaghat S, et al. Building cancer registries in a lower resource setting: the 10-year experience of Golestan, Northern Iran. *Cancer Epidemiol.* 2018;52:128-33. doi: 10.1016/j.canep.2017.12.014.
  16. Mendy M, Caboux E, Lawlor RT, Wright J, Wild CP. Common Minimum Technical Standards and Protocols for Biobanks Dedicated to Cancer Research (IARC Technical Publications; 44). Lyon, France: International Agency for Research on Cancer; 2017.
  17. Campbell LD, Betsou F, Garcia DL, Giri JG, Pitt KE, Pugh RS, et al. Development of the ISBER best practices for repositories: collection, storage, retrieval and distribution of biological materials for research. *Biopreserv Biobank.* 2012;10(2):232-3. doi: 10.1089/bio.2012.1025.
  18. Fritz AG. International Classification of Diseases for Oncology: ICD-O. World Health Organization; 2000.
  19. Pourshams A, Khademi H, Malekshah AF, Islami F, Nouraei M, Sadjadi AR, et al. Cohort profile: the Golestan Cohort Study--a prospective study of oesophageal cancer in northern Iran. *Int J Epidemiol.* 2010;39(1):52-9. doi: 10.1093/ije/dyp161.
  20. Spencer LT, Bancroft JD. Tissue processing. In: Suvarna SK, Layton C, Bancroft JD, eds. *Bancroft's Theory and Practice of Histological Techniques.* 7th ed. Oxford: Churchill Livingstone; 2013. p. 105-23.
  21. Ervik MJ, Cooke AP, Ferlay J, Rahimi A, Antomi S, Dhivar D, et al. CanReg5: Computer Software for Cancer Registries. Lyon: International Agency for Research on Cancer; 2008.
  22. Cancer Research UK. Grand Challenge Mutographs project. Available from: <https://www.mutographs.org/mutographs-project/partner-institutes/>. Accessed November 14, 2020.
  23. Biobank and Cohort Building Network (BCNet) Lyon: International Agency for Research on Cancer; 2020. Available from: <https://bcnet.iarc.fr/>. Accessed December 8, 2020.
  24. Ghasemi-Kebria F, Amiriani T, Fazel A, Naimi-Tabiei M, Norouzi A, Khoshnia M, et al. Trends in the incidence of stomach cancer in Golestan province, a high-risk area in northern Iran, 2004-2016. *Arch Iran Med.* 2020;23(6):362-8. doi: 10.34172/aim.2020.28.
  25. Roshandel G, Ferlay J, Semnani S, Fazel A, Naeimi-Tabiei M, Ashaari M, et al. Recent cancer incidence trends and short-term predictions in Golestan, Iran 2004-2025. *Cancer Epidemiol.* 2020;67:101728. doi: 10.1016/j.canep.2020.101728.
  26. International Agency for Research on Cancer. BCNetter, October 2019 Lyon: International Agency for Research on Cancer. Available from: <https://bcnet.iarc.fr/docs/BCNetter10-2019.pdf>. Accessed November 14, 2020.