The foundation of personalized medicine is the establishment of biobanks and their standardization

Liliana Policiuc[1,2]*, Mihail Buse[3]*, Diana Gulei[3], Alina Andreea Zimta[3], Cornelia Braicu[1], Alexandra Iulia Irimie[4], Ioana Berindan-Neagoe[1,2,3]

[1]Research Center for Functional Genomics and Translational Medicine, "Iuliu Hatieganu" University of Medicine and Pharmacy, Cluj-Napoca, Romania; [2]Department of Functional Genomics and Experimental Pathology, "Prof. Dr. Ion Chiricuta" Oncology Institute, Cluj-Napoca, Romania; [3]MEDFUTURE -Research Center for Advanced Medicine, "Iuliu Hatieganu" University of Medicine and Pharmacy, Cluj-Napoca, Romania; [4]Department of Prosthodontics and Dental Materials, Faculty of Dental Medicine, "Iuliu Hatieganu" University of Medicine and Pharmacy, Cluj-Napoca, Romania

*These authors contributed equally to this article

Summary

The present review outlines the current information available about biobanks including: summarizing the main ethical issues encountered when integrating patient samples in research projects; emphasizing the importance of biobanks as the basis for any personalized medicine therapies in cancer with a particular focus for oral cancer, through the use of biomarkers; and providing examples of biobanks that use anonymous sample collection and labelling methodologies to help alleviate the problems of privacy, informed consent, data security as well as public trust.

Key words: *biobanks, cancer, personalized medicine*

Introduction

The discovery and rapid development of personalized medicine is now dominating the diagnostic, prognostic and therapeutic approaches taken toward a variety of diseases including cancer. This emphasis on patient genetic and molecular profiles stimulates an ever increasing demand for biobanks. The development of these personalized therapies toward disease incorporates the organization of patients according to their genetic profile, lifestyle, risk factors and demographic characteristics which are inherent to biobanks; not to mention that biobanks can provide the physical samples essential for experimentation. In particular, oral cancer could be an essential contributor to the establishment of biobanks because of the non-invasive samples that can be obtained from patients [1]. The genetic material obtained from these samples could reveal the genes responsible for the activation of early carcinogenic mechanisms, disease progression or metastasis [2]. Therefore, oral cancer can facilitate the development of a representative disease model using biobanks that could be applied to all cancers.

More than 90% of the patient cases with oral cancers are oral squamous cell carcinoma (OSCC) which occur on the epithelial lining of the oral cavity [3]. Due to the non-invasive approach of obtaining saliva biosamples, patients would more readily

donate, thereby, intensifying the necessity for biobanks because of their storage capabilities. It is important to note that the lack of standardization for collection and processing protocol exists for all biosamples. The differences in the factors related to collection and processes raises the question of whether salivary biomarkers can be compared between labs [3], thereby, oral cancer could provide the initial step toward implementing a standardized biobanking system, concurrently resolving ethical, sample collection and processing protocols.

Genetic variation, with regard to personalized medicine, regulates disease susceptibility and drug response. Monitoring these genetic variations in both healthy individuals and patients can be accomplished through biobanking storage of biomarkers. All of this hinges on the public's trust and willingness to donate their biosamples. The focus of the public's attention shifted from the importance of the establishment of biobanks to the ethical issues that have arisen from un-standardized informed consent and more specifically, the grade of restriction that the informed consent imposes on the patient samples toward the research in a particular project or further experimentation in future projects. Therefore, the implementation of an ethical framework cannot be delayed. Science and ethics need to progress in accord not only to gain the public's trust but also to ensure proper implementation of advanced research projects.

The biobank, in its simplest and most encompassing definition, is a repository for the organized collection of biological samples associated to personal health information that will be used for biomedical research [4,5]. The Pan-European Biobanking and Biomolecular Resources Research Infrastructure (BBMRI) grouped biobanks into two categories: population or disease-based. Population biobanks are concerned with the evolution of different conditions over time while disease-based biobanks target specific diseases from the patient sample in order to develop personalized forms of treatment [6]. These biobanks are best utilized together; for example, population-based biobanks depend on the results of clinical biobanks for the precise characterization of phenotype and molecular profiles, while clinical biobanks need biological matter collected at earlier time points from the population-based cohorts for controls [7]. The increasing number of biobanks in oral health and personalized dentistry demonstrate the impact that biobanks have in this post-genomic era [8-11]. One example is the Malaysian Oral Cancer Database and Tissue Bank System (MOCDTBS) that preformed a compressive clinical data collection with emphasis on the environmental factors [12].

Ethics

The concept of informed consent was first established by the Nuremberg Code which necessitated the voluntary consent of the human subject in a full legal capacity. Furthermore, well-informed understanding of the research should be given such as: the nature, duration and purpose of the experiment; the methods and means to which it was conducted; all reasonable risks and hazards to be expected; and the health benefits for the patient is possible from participation in the experiment [13]. These requirements have been modified since then into the Declaration of Helsinki; the latest of them was in 2013, adding that research participants should know the sources of funding, possible conflict of interests, intended benefits and, most importantly, the right to withdraw. Furthermore, specifically relating to biobanks or similar repositories, the physician must seek informed consent for the collection, storage and use of identifiable human material or data [4,13]. For biobanking, these standards of informed consent can be extremely problematic for four reasons: firstly, informed consent only applies to the participant and genetic information is connected to all the participant's family/relatives; secondly, due to the nature of research, consent cannot always be informed at the time since the future of the project is not known; thirdly, biobanks are a resource for research projects and do not in fact carry out the research; fourth, the "right to withdraw" may not always be possible for biobanks [9,13].

Consent is the main and largest issue with regard to the establishment of a biobank. Consent is the participants' right to decide whether donation from their body in the form of a biospecimen can be used for research purposes. All authors and researchers strongly agree that participants without consent should not donate or come to any risk from the research being done. Furthermore, they also agree that consent should be done in an informed manner such that the participants understand the purpose of the research, expected benefits and risks associated. To ensure additional protection, respect and trust, a majority of authors support the notion of withdrawing consent [5]. Therefore, the issues surrounding consent do not revolve around intention but rather the type of consent given. There are four types of consent that can be given: firstly, broad consent which is a general form giving researchers the ability to freely use the collected samples in any research project without necessitating the new consent for a future or new project; secondly, partially restrictive consent which means that the participant allows the

collected samples to be used in any studies that are connected to the one they initially gave compliance; thirdly, multi-layered consent gives the participant options explained by the researcher in detail; lastly, specific informed consent which allows the participants' samples to only be used in the particular study and does not allow for the sample to be used in any other study or context. One distinction that needs to be made is that broad consent does not mean blanket consent to all uses. In agreement with legislation in many countries, consent refers only to use in biomedical research, for example not commercial or forensic practices [14].

There is no consensus between researchers over which is the best type of consent to be given. Some argue for a standardized universal consent to ensure repeatability of protocol and comparable results, while others feel that this type of consent does not take into account geographical, social and religious differences of the participants which must be taken into consideration [5]. One method of taking all of this into consideration would be to establish an ethical review board as a separate entity from the biobank and use it to govern the research being done. Ethical review board would be formed locally, meaning that they would take into account geographical, social and religious differences of the participants. Furthermore, these ethical review boards provide the legal frameworks for the biobanks, thereby, ensuring the appropriate use of the samples and participant protection. By doing all this and integrating the public into the research decision-making process, public trust would be maintained and encouraged [5,15].

Some researchers argue that there is no informed consent that can be given because research by nature is specific and that there is not enough information to fully inform the participants [13]. Generally, research is aimed at answering an unknown question, or phenomena, in a specific field and this means that at the time of consent there is little information to inform the participant about risk. The Declaration of Helsinki assumes that consent takes place at the beginning of the research where the future aims, benefits and risks are known. However, this is not the case with biobanking where the goal is to acquire large collection of biospecimen samples representative of a population for un-specified research [14], such that research institutions can later draw from this bank for accurate high-quality samples for experimentation. It is true that research institutions can establish biobanks prior to the project or as a part of the project itself, however, the goal remains the same. The main argument would be that biobanks provide the way consent can be informed. One

example is the UK biobank which specifically states that it is a resource for supporting diverse range of research intended to improve prevention, diagnosis and treatment of illness, thereby, promoting health throughout society. It is impossible to predict how the research will progress and what discoveries will be made, which can then modify the research objective from what it was initially. In conclusion, the future-oriented nature of biobanks makes it impossible for consent in biobanking to ever be fully informed. Consent lies with the research institution because biobanks just store information like a library. Research institutions use biobanks, meaning that the consent should be in terms of the research and how the samples will be used [13].

Two practical problems that have arisen regarding consent from biobanks under the Declaration of Helsinki are: consent must be given for every single sample from participant and the right to withdraw consent at any time. Due to the fact that biobanks aim at containing large sample sizes on the scale of populations it is administratively impractical and unrealistic to get consent for every individual for every research project; this becomes impossible when you take into consideration the notion of re-consent. Campbell (2007, p.242) goes as far as saying that to truly promote trust and altruism between the public and biobanks, one should refrain from suggesting continual consent over the use of donated sample, meaning that re-consent inquiries raise more ethical questions about the use of the biospecimen [16]. Furthermore, the right to withdraw consent at any point may be impossible. If the biospecimen has been anonymized and tested upon, there is no way to remove it from the sample collective. The best case scenario would be that future research projects could not use the biospecimen retrieved from the participant because of its removal from the database and its discarded from the biorepository. Therefore, the only solution is to inform the participants of a time limit they have to withdraw consent from the current research. The declaration of Helsinki does stipulate that in exceptional situations where consent is impossible or impractical the research ethical committee would give final approval, meaning that the right to withdraw could also fall under this authority.

The last essential issue that needs to be addressed, specifically by genetic biobanks, is privacy. Simply put, the genetic profile provides information about the family related to the participant. This makes the individual's informed consent not encompass the privacy of third parties. This leaves the possibility of discrimination or misuse of

information based on cultural, geographical and age-related groups [13]. Furthermore, due to the fact that genetic information contains more than just the information that the certain research project is purposed with studying, the rest of the genetic information can be accessed at any time for purposes beyond what the participant initially informed consent. In April 2011, Arizona State University had a law suit filed against them from the Havasupai Indian tribe of Arizona for the misuse of their genetic information from blood samples. The researchers' broadly worded consent document such that the tribe donated for diabetes research, when in reality the researchers used the information to study the genetics of mental illness ancestry [17,18]. This problem here lies with identification of individual and relatives from the genetic information obtained. The identity needs to be removed while still maintaining the important biological characteristics found within the genetic information. This process of de-identification of biospecimens stored in a biobank is known as anonymization. Anonymization must be handled with care because the potential of a biobank lies in the ability to correlate genetic or biological data to phenotypes, medical and personal information.

Biobanks begin and end with the participant donating biospecimens to the research because without this donation, the research cannot progress and the biobanks would cease to exist. It is in the biobanks best interest to constantly improve practices in the ethical, legal and social fields in three ways: protection of rights and welfare of participants, demonstration of participant respect and promotion of ethically responsible research practices [19].

Personalized medicine & biomarkers

The long-standing approach to the treatment of diseases in medicine has been a conservative "one size fits all" model, focusing on fitting patients' symptoms into a disease type and treating the patient with the statistically most successful therapy for the disease-type population [20]. Conversely, personalized medicine allows for the possibility for an accurate and individualized diagnosis that contributes to a precise and targeted treatment in order to reduce negative outcomes or side effects [21]. The rapid evolution of technologies that allows scientists to acquire and analyze high-dimensional data like genomes or proteomes
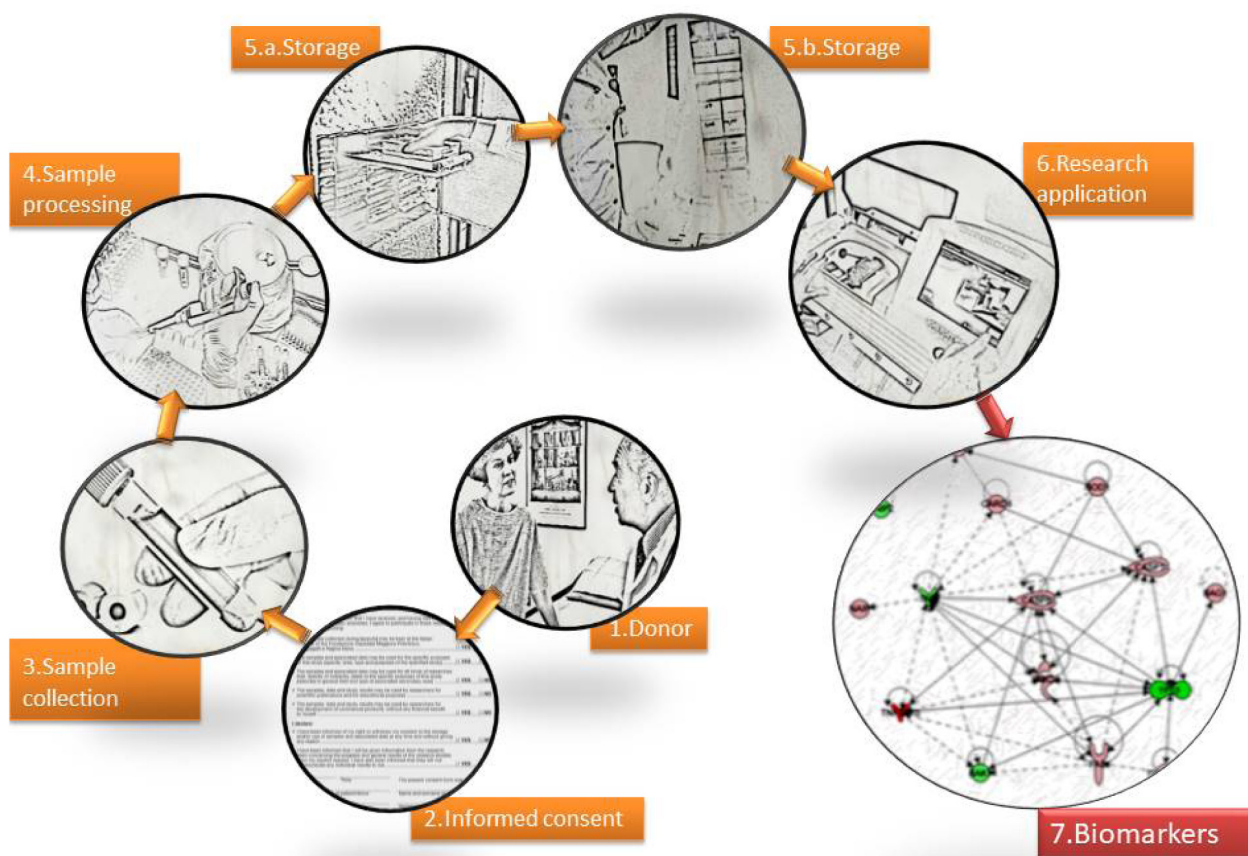

Figure 1. Flowchart of the biobank system's process starting with clinical evaluation of donor patient till the research of developing biomarkers for prognosis or diagnosis.

created this potential for personalized medicine [21]. In this respect, the most recent promising fields of research towards linking genotype-phenotype are bionetwork modeling [22]. Furthermore, the challenge lies in very large sample sets not just for discovery but, more importantly, for validation. This provides the rational for establishing biobanks. Figure 1 depicts the biobanking procedure starting from participant consent and sample collection, through obtaining biomarkers from genetic profiling techniques.

Personalized medicine can then improve the prevention and cure of the disease by predicting both the disease risk within a general healthy population and the therapeutic response within patients [11,23,24]. This prediction comes in the form of biomarkers, which are derived from patient genomes. According to the Biomarkers Definitions Working Group (2001), a biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biological process, pathogenic process or pharmacologic responses

to a therapeutic intervention [25]. Simply put, biomarkers provide genetic information encoded within DNA. DNA is stable over an entire life time, making these biomarkers easily measureable at any point in time; this facilitates the scientific methodology of repeatability and reproducibility. Furthermore, this means that they can be used in both prospective and retrospective studies. The biomarkers that are deemed stable are termed DNA biomarkers and they include DNA sequences such as single nucleotide polymorphisms (SNPs) or short tandem repeats (STRs). Currently, due to the availability of high-quality molecular profiling techniques and facilities, SNPs are the most common used type of DNA variant. It was found that in most situations SNPs have two different alleles at a particular gene locus which results in three possible genotypes within the patient. Originating from the fact that cancer is a disease that changes DNA on the cellular level, biomarkers can be used to measure these changes in tumor [26]. However, DNA biomarkers are not without disadvantages.
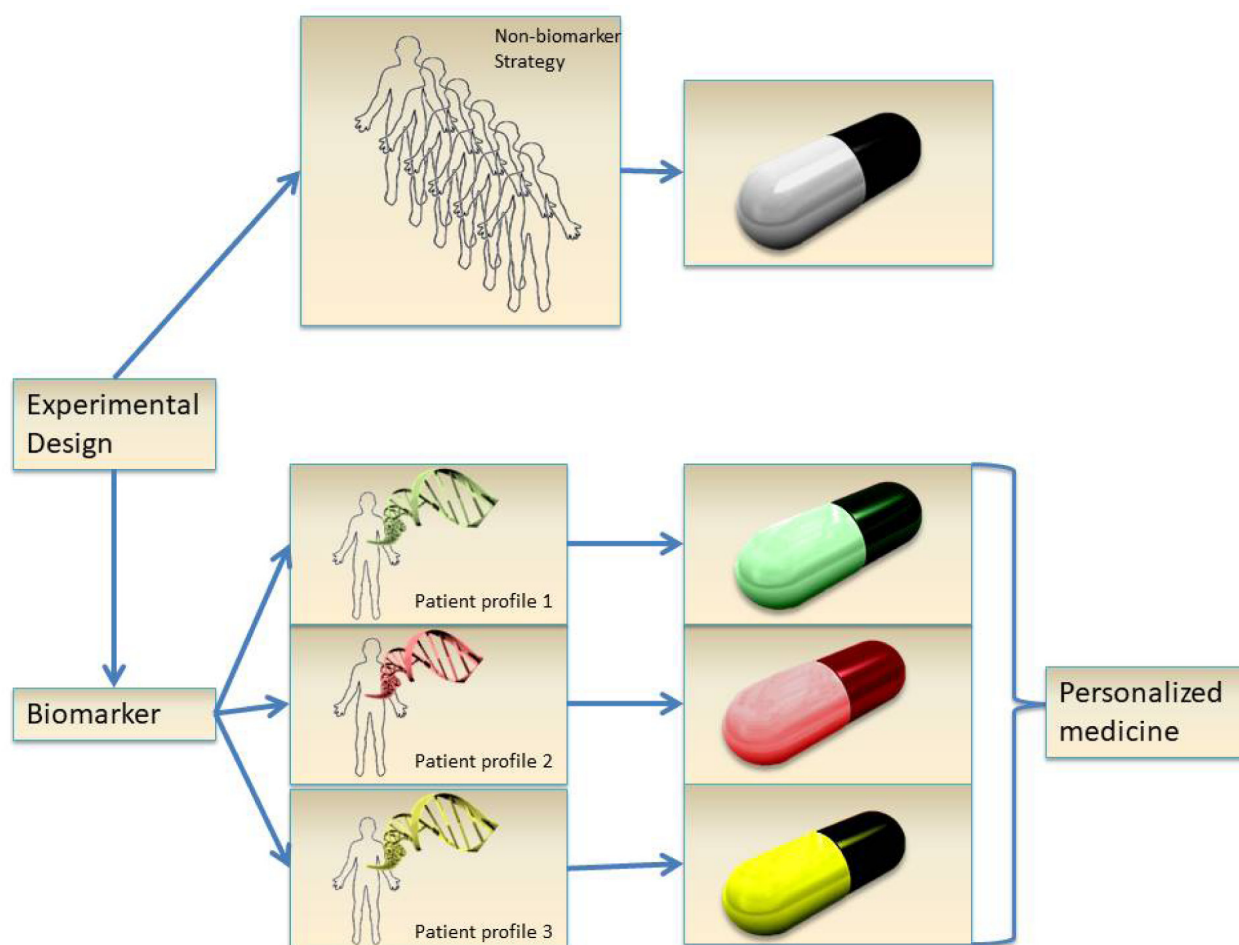

Figure 2. Representation of achieving a personalized therapy by using biomarkers and treatment drug responses. The experimental design is based on a chosen biomarker of research interest and by applying single different treatments to each patient. If the treatment is successful then a desired outcome has been seen in the biomarker, meaning successful treatments can then be combined to create a personalized therapy.

The discovery rate of new DNA biomarkers is faster than their product cycles. This means that by the time a DNA marker test has passed all the steps for marketing approval, it may have already been rendered scientifically less significant than a newly discovered DNA biomarker with better performance [26].

Nonetheless, DNA biomarkers have an important role to play with biobanks. The biomarker-defined population subgroups used for systematic investigations into the improved efficacy and safety of approved drugs require validation that can only be achieved using biobanks. Biobanks contain large repositories of biospecimens that can be used for genotyping, thereby, facilitating the validation process. As Wang et al. stated in 2010, valid conclusions from genotyping cannot be drawn from only from convenience samples or small sample sizes [27]. The individualized profile design proposed by personalized medicine requires an accumulation of a large number of different patient profiles which, through a comparative validation strategy, can lead to the selection of a treatment from a large number of different therapies comparing the conventional treatment selection to an individualized treatment decision. Such a treatment plan necessitates regulation of the particular application of drugs for a treatment, substantiated by clinical trials data. For the patient to get best result during complex decision-making process, validation is necessary for both the biomarker-based treatment selection and the effectiveness of the individual treatments. This can allow for several monotherapies, each selected on the presence or absence of specific DNA variants (biomarkers), to be combined for an individualized therapy [26]; this process can be seen in Figure 2.

Within personalized medicine another division is made between biomarkers, i.e. they can be diagnostic, prognostic or preventive. The difference lies in the fact that diagnostic and prognostic biomarkers help predict disease progress while predictive biomarkers are determined in response to a treatment [26,28]. The establishment of biobanks could help in three main ways: determining genetic markers of the most clinical significance; limiting off-target effects of gene-based therapies; and conducting clinical studies to identify the genetic variants correlated to drug response [29]. Moreover, studies identifying many genetic variations
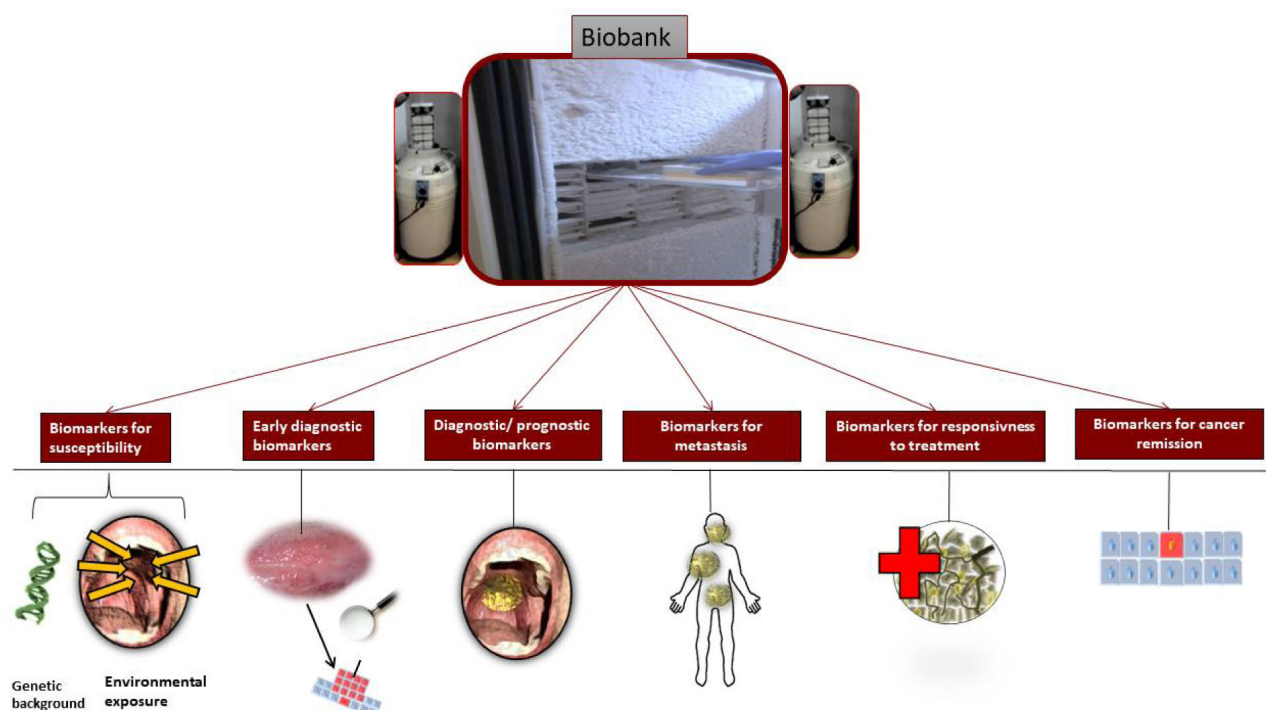

Figure 3. The biomarkers are snapshots of a disease progression, from its basis, the inherited genetic predisposition that a person is born with (biomarkers for susceptibility), which combines with the environmental factors that affect the human organism (biomarkers for exposure), leading to the installment of early biological lesions (early diagnostic biomarkers), disease progression to hyperplasia (hyperplastic biomarkers) and ultimately to cancer (diagnostic and/or prognostic biomarkers). From this point on, the disease can run its course and become metastatic (metastasis biomarkers), or the patient can receive different treatment options (treatment responsiveness biomarkers) and enter remission (remission biomarkers) or the cancer can appear again (recurrence biomarkers). With the help of biomarkers the targeted therapy can be developed, a better option for cancer patients that has better chances of ending disease progression.

underlying risks of rare or common diseases facilitate the drug targeting of genes, proteins or pathways [30]. Figure 3 shows how the progression of cancer relates to the types of biomarkers used in personalized medicine, specifically indicating which kind of biobank stores each biomarker [31].

Moreover, diagnostic biomarkers are used to determine the severity of a disease, i.e. to screen for the differences between healthy individuals and those at an early stage of disease. For example, the point-of-care tests Rheuma-Chec and CCPoint screen for rheumatoid arthritis in non-symptomatic, healthy individuals by testing their serum for mutated citrullinated vitmentin and citrullinated peptides/proteins antibodies [32]. Prognostic biomarkers are used for the prediction of a disease in a defined clinical population under standard treatment conditions. MammaPrint is a DNA tumor biomarker for breast cancer prognosis post-surgery to determine the risk of metastasis. It correctly predicted breast cancer prognosis to the point that the FDA approved its use as an *in vitro* diagnostic multivariate index assay [33]. Predictive biomarkers are used to predict the response of a patient to a treatment in terms of drug effectiveness and safety; this prediction is an important factor in clinical decision making. In 2011 Keedy et al. used the epidermal growth factor receptor mutation in patients with advanced non-small cell lung cancer to determine whether patients were eligible for the first-line EGR tyrosine kinase inhibitor therapy [34].

With regard to cancer research, one of the challenges with exponentially increasing urgency is the availability of adequately annotated and appropriately collected biospecimens. Frozen and formalin-fixed paraffin-embedded (FFPE) tissue formats are commonly used to preserve biospecimens for future use. A study done in 2010 by Hughes et al., analyzing four major cancer research journals (Cancer Research, Clinical Cancer Research, British Journal of Cancer and International Journal of Cancer) found a threefold increase in the average cohort size for both FFPE and frozen tissue biospecimens over the past 20 years [35]. Secondly, the quality of the biosample can be ensured with the standardization of biobanks. Microarray gene expression profiling has become an essential technique allowing for the identification of molecular cancer subtypes, capturing the heterogeneity of cancer. The quality of the gene expression is definitively coupled with the quality of the biospecimen and the degree of degradation from the extracted RNA. Biospecimen quality can be affected by *in vivo* ischemia time, surgical manipulation and anesthesia during the process of specimen retrieval [36,37].

In oral cancer a particular focus was on salivary biomarkers that represent a promising non-invasive approach and a domain of strong research interest [38]. The main issue related with these types of biological fluid has to do with the missing of standardized approach for collection, processing, and storage [38]. This is related with a high variability among the tested biomarkers in different patient cohorts and can be avoided by the development of biobanks [3,38].

Importance of oral cancer research to biobanking

Human biological specimens in oral cancer, that include tissue (normal/tumoral), whole blood, serum and saliva as particularity for oral cancer were widely used for translational research to understand the disease pathogenesis and to evaluated different scientific hypotheses in the context of biomarker discovery. The success of these studies is affected by the quality of the material used. Thus, the aforementioned problems regarding sample quality would be resolved with a standardized biobanking system. This would give better molecular profiles which are in dire need for cancer research to further develop the concept of personalized medicine. However, not all patient samples need to be invasive or require a clinical procedure to remove them. In recent years, more and more non-invasive types of patient samples are proving to be valuable and credible with regard to their molecular profiles. Saliva has become a distinctive diagnostic fluid because of its simple collection, processing and minimally invasiveness [39]. Whole saliva is unique and complex regarding both composition and source. The majority of saliva consists of proteins and polypeptides; specifically, more than 2300 proteins and peptides have been found in human saliva [3] . However, the most abundant, which comprise 98% of the total proteins, are: α-amylase, albumin, cystatins, hystatins, secretory-IgA, lactoferrin, mucins, lysozymes, proline-rich proteins, statherin and transferrin. Thus, the remaining proteins are at very low concentrations meaning that saliva analysis for possible biomarkers requires extremely sensitive methods or instruments for detection [3]. The largest amount of research being done on saliva and its potential to reveal biomarkers has been for oral cancer.

Multiple studies have found useful biomarkers in saliva obtained from oral cancer patients [23,24,40,41]. One of the first studies done to demonstrate the diagnostic value of RNA found in human saliva was by Li et al. [42], utilizing OSCC as the proof-of-principle disease. They performed a

qPCR and compared the saliva of OSCC patients to that of controls, discovering that seven salivary mRNAs had at least a 3.5-fold elevation. The transcripts of these mRNAs could act as biomarkers and they include *IL8, IL1B, DUSP1, HA3, OAZ1, S100P* and *SAT*. The combination of these biomarkers yielded a high sensitivity and specificity in distinguishing OSCC from the controls. Therefore, salivary transcriptome profiling could prove to be useful in other major diseases as well [42]. Similar successful discoveries of oral fluid biomarkers were validated by mass spectrometry [43] and by subtractive proteomics which consisted of shotgun proteome analysis, 2D-gel electrophoresis and immunoassays [43]. In 2010, another group of scientist identified salivary transferrin as a biomarker for early detection of oral cancer. They used a combination of all the previous mentioned techniques, i.e. 2D-gel electrophoresis and MALDI-TOF mass spectrometry to analyze the protein pool of both OSCC and OSCC-free persons, followed by western blot and ELISA assays for verification of potential biomarkers [44]. Lastly, one study distinguished that the inflammatory cytokine proteins interleukin-8 and interleukin-1beta as having significantly higher levels in patients with OSCC when compared to healthy control subjects. This was done by implementing luminex xMAP single-plex and multi-plex assays, which proved to be just as effective as ELISA assays for the quantification of proteins in saliva [39]. The only criticism for the majority of these studies was that their sample sizes were too small to extrapolate for populations. Biobanks could provide databases and biorepositories, such that individual studies could be combined with population-based results. When analyzing saliva samples from OSCC patients from a Serbian population, three proteins and four miRNAs were validated as biomarkers; these seven biomarkers performed strongest during the late stage of the cancer. The salivary protein biomarkers were interleukin-1beta, interleukin-8 and s90K/Mac-2 binding protein (M2BP) while the salivary RNAs markers were SAT1, S100P, IL8 and IL1B. According to the authors of this article, this also coincides with the salivary biomarkers found in OSCC patient population in the USA [45].

Biobanking – Recent approaches

One example of a successful cancer case-control study that used Nordic biobanks and cancer registries was done in 2008 by Kapeu and his colleagues investigating whether smoking is an independent risk factor for invasive cervical cancer [46]. This study used five population-based serum banks: Finnish Maternity Cohort, Icelandic Maternity Cohort, Northern Sweden Maternity Cohort, Northern Sweden Health and Disease Study (comprising of MONICA project and Vasterbotten Intervention Program) and Janus Serum Bank. These serum banks which contained samples from more than 1,000,000 subjects were linked to national cancer registries of Finland, Iceland, Norway and Sweden. Specifically, in this study 588 samples of female serum samples were analyzed who developed invasive cervical cancer, and were compared to 2,861 matched controls. They used a tobacco exposure biomarker cotinine as well as antibodies from HPV (types 16 and 18), herpes simplex virus type 2 and *Chlamydia trachomatis*. The first major finding was that smoking was associated with increased risk of squamous cell carcinoma/cervical cancer after adjusting antibodies to oncogenic HPV; specifically, there was an increased 2-fold excess risk of squamous cell carcinoma among HPV16- and HPV18-seropositive heavy smokers. The second major finding was that the risk of squamous cell carcinoma increased with increasing cotinine dose and age at diagnosis [46]. These authors, supported by the studies of Suk et al. (2001) and Poppe et al. (1995), believe that this is due to the fact that cigarette smoke enhances the carcinogenic potential of HPV by: firstly, the inhibition of apoptosis through targeting interferon-gamma and tumor necrosis factor-alpha; and, secondly, enhancing HPV ability to evade the immune system through the manipulation of cytokine expression [47,48]. This Nordic joint study is the largest conducted to date on the risk factors of cervical cancer. If it were not for the large sample size and human serum obtained from the various biobanks, such clear results could not be drawn. This study provides a milestone and guide for the correct use of biobanks as well as the ethical cooperation between all the parties. For oral cancer an important source of biomarkers is represented by whole saliva, that contains fluids produced by salivary glands; saliva comprises a wide range of gingival fluids and cellular debits with important diagnostic and prognostic value [49].

With the increasing demand for biobanking and the computational technologies allowing for electronically advanced protection, some anonymous or coded systems for sample registration, collection and labeling have been developed to bypass some of the ethical issues. Furthermore, there is a shift occurring to rely on electronic medical records to produce self-contained systems for the integration of genetics and medical information. The only drawback is that the process of converting all this information into an electronic format

is very time-consuming and costly [18]. This all falls under the concept of de-identifying the data from the participant while still remaining with the biological information pertinent to the research.

To date there are four of these kind of systems: Vanderbilt BioVU DNA databank, deCODE genetics, Netherlands Cancer Institute and Bio-PIN. Bio-PIN and Vanderbilt BioVU DNA databank use an anonymous registration [21] while deCODE genetics and Netherlands Cancer Institute use a coded registration [50,51]. Bio-PIN and deCODE genetics allow for the combining data between different institutes, such that research projects can have large sample sizes and information at their availability [50,52]. However, only Bio-PIN actually allows for two-communication between the participant and the biobank database on a web-based platform.

Bio-PIN

Bio-PIN is a biobanking system proposed in 2011 by Neitfield et al., wherein samples and associated data would be deposited anonymously and labeled using a PIN code that is generated, based on the biological characteristics of the patient [52]. This concept is based on the idea that each individual patient has unique biological characteristics derived from every biological sample which can then be transformed into a simple alphanumerical code. Proposed examples of distinguishing biological characteristics (DBC) could be SNPs, short tandem repeats or biochemical markers. Therefore, the Bio-PINs contains no identity data or individually identifiable data making it anonymous according to delineations expressed by HIPAA of 1996. Important to note that the probability of finding identical distinguishing biological characteristics is not zero; to address this, a birth order number is also added to the DBC which would result in different Bio-PIN in all cases. The Bio-PIN would be created in such a way that it could not be de-transformed back into the original DBC. Using Bio-PINs for biobanking can be summarized into four steps: firstly, the anonymous collection and registration of biological samples in the biorepository along with associated data in the biobank database; secondly, the determination of the distinguishing biological characteristic from each sample followed by, subsequent, generation of the Bio-PIN; thirdly, connecting the Bio-PIN to its respective sample in the biorepository records to the database; finally, establishing a communication with the individual patient only using their respective Bio-PIN [52]. The communication between the patient and the biobank using the Bio-PIN is essential to encourage participants' regular access to their personal record in the biobank database using a web-based platform, not only for verification or updates, but this could also be the platform from which they can inform themselves on the research projects and provide anonymous consent to the research projects. With these consent proposals, additional information could be asked like lifestyle, environmental and medical history. PatientsLikeMe and 23andMe are two biobanking companies that successfully implemented a website to integrate with the general public. PatientsLikeMe has built an extensive repository of detailed phenotypic ad health data that is public-sourced from a community of over 115,000 participants. 23andMe contains biospecimen samples and genomic data from 120,000 participants, of which 87% opted to donate to research. These web platforms allow participants to control the information they provide, see how it is used and are regularly updated on new discoveries. The result is that participants have chosen to give more than the minimum, demonstrating their trust and enthusiasm [53,54]. In essence, the more regularly the participant deposits accurate and complete data, the more constructive their samples would be to the research and thereby produce additional diagnostic or prognostic information important to their health. For example, these results could allow for testing diagnostic or prognostic values for life-threatening diseases. This would then become available to the participant under categories ranging from 'presently little health risk' to 'seeking medical advice is strongly advised.

In conclusion, the most significant advantage of the Bio-PIN system is that it requires a specific consent from the participants that can be modified at any time. The choice of consent is dependent on the level of participation; 'active' participants can continuously exercise their autonomy and consent while 'inactive' participants *de facto* give broader consent to the available research initiatives obtained during the registration. This participation could also extend to medical institutions where any biological samples leftover from Bio-PIN participants could be transferred as anonymous material to the biobank, rather than being wastefully discarded [52].

Conclusions

This review had the purpose to present in an objective approach the current situations regarding the biobanking activity. The first steps need to be made in the establishment of biobanks such that standardization can occur and ethical issues can be resolved. Standardization of practices has shown to improve the quality of the biospecimen, thereby

improving the accuracy of the results. Implementation of the Bio-PIN biobanking method, would inherently lead to improvements in registration, archiving, tracing and sample usage but also the legal and ethical issues regarding biobanks. Despite the rapid advancements in the research using biobanks, we are currently being held behind by the ethical aspects concerning the usage of the donated samples. Instead of making the research institute accountable for consent of participant, we are forcing biobanks to address the issue despite the fact that they are just a resource for research. Furthermore, biobanks provide a way for research-ers to become informed such that they can give the patients all the best form of consent. Despite all this, there are collaborations between research institutes and biobanks that are doing fundamental work to identify genetic variants underlying phenotype that can lead to personalized medicine. The implementation of targeted and personalized therapies toward disease management facilitates the translational relevance toward clinical practice.

Conflict of interests

The authors declare no conflict of interests.

References

1. Psoter WJ, Morse DE, Sánchez-Ayendez M et al. Increasing opportunistic oral cancer screening examinations: findings from focus groups with general dentists in Puerto Rico. J Cancer Education 2015;30277-83.

2. Vijay Kumar KV, Suresan V. Knowledge, attitude and screening practices of general dentists concerning oral cancer in Bangalore city. Indian J Cancer 2012;49:33-8.

3. Cheng YS, Rees T, Wright J. A review of research on salivary biomarkers for oral cancer detection. Clin Transl Med 2014;3:3.

4. Cambon-Thomsen A, Rial-Sebbag E, Knoppers BM. Trends in ethical and legal frameworks for the use of human biobanks. Eur Respir J 2007;30:373-82.

5. Budimir D, Polasek O, Marusic A et al. Ethical aspects of human biobanks: a systematic review. Croat Med J 2011;52:262-79.

6. Asslaber M, Zatloukal K. Biobanks: transnational, European and global networks. Brief Funct Genomics Proteomics 2007;6:193-201.

7. Gottweis H, Zatloukal K. Biobank governance: trends and perspectives. Pathobiol J 2007;74:206-11.

8. Birch K, Dove ES, Chiappetta M, Gursoy UK. Biobanks in Oral Health: Promises and Implications of Post-Neoliberal Science and Innovation. Omics 2016;20:36-41.

9. Braicu C, Berindan-Neagoe I, Pileczki V et al. Breast tumor bank: an important resource for developing translational cancer research in Romania. Cancer Biomark 2014;14:119-27.

10. Ionescu C, Braicu C, Chiorean R et al. TIMP-1 expression in human colorectal cancer is associated with SMAD3 gene expression levels: a pilot study. J Gastrointest Liver Dis 2014;23:413-8.

11. Braicu C, Tudoran O, Balacescu L et al. The significance of PDGF expression in serum of colorectal carcinoma patients--correlation with Duke's classification. Can PDGF become a potential biomarker? Chirurgia 2013;108:849-54.

12. Zain RB, Athirajan V, Ghani WM et al. An oral cancer biobank initiative: a platform for multidisciplinary research in a developing country. Cell Tissue Banking 2013;14:45-52.

13. Widdows H, Cordell S. The ethics of biobanking: key issues and controversies. Health Care Analysis 2011;19:207-19.

14. Hansson MG. Ethics and biobanks. Br J Cancer 2009;100:8-12.

15. De Vries RG, Tomlinson T, Kim HM et al. The moral concerns of biobank donors: the effect of non-welfare interests on willingness to donate. Life Sci Soc Policy 2016;12:3.

16. Campbell AV. The Ethical Challenges of Genetic Databases: Safeguarding Altruism and Trust. King's Law J 2007;18:227-45.

17. Mello MM, Wolf LE. The Havasupai Indian tribe case-lessons for research involving stored biologic samples. New Engl J Med 2010;363:204-7.

18. Scott CT, Caulfield T, Borgelt E, Illes J. Personal medicine - the new banking crisis. Nat Biotechnol 2012;30:141-7.

19. Vaught J, Lockhart NC. The evolution of biobanking best practices. Clin Chim Acta 2012;413:1569-75.

20. Kinkorova J. Biobanks in the era of personalized medicine: objectives, challenges, and innovation: Overview. EPMA J 2015;7:4.

21. Roden DM, Pulley JM, Basford MA et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. Clin Pharmacol Ther 2008;84:362-9.

22. De Palma M, Hanaan D. Individual genomes and personalized medicine: life diversity and complexity. Personal Med 2010;7:347-50.

23. Braicu C, Cojocneanu-Petric R, Chira S et al. Clinical and pathological implications of miRNA in bladder cancer. Int J Nanomed 2015;10:791-800.

24. Irimie AI, Sonea L, Jurj A et al. Future trends and emerging issues for nanodelivery systems in oral and oropharyngeal cancer. Int J Nanomed 2017;12:4593-4606.

25. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 2001;69:89-95.

26. Ziegler A, Koch A, Krockenberger K, Grosshennig A. Personalized medicine using DNA biomarkers: a review. Human Gen 2012;131:1627-38.

27. Wang SJ, O'Neill RT, Hung HJ. Statistical considerations in evaluating pharmacogenomics-based clinical effect for confirmatory trials. Clin Trials 2010;7:525-36.

28. Buyse M, Michiels S, Sargent DJ, Grothey A, Matheson A, de Gramont A. Integrating biomarkers in clinical trials. Expert Rev Molec Diagn 2011;11:171-82.

29. Hamburg MA, Collins FS. The path to personalized medicine. New Engl J Med 2010;363:301-4.

30. Watson RW, Kay EW, Smith D. Integrating biobanks: addressing the practical and ethical issues to deliver a valuable tool for cancer research. Nat Rev Cancer 2010;10:646-51.

31. Riegman PH, Morente MM, Betsou F, de Blasio P, Geary P, Marble Arch International Working Group on Biobanking for Biomedical Research. Biobanking for better healthcare. Molec Oncol 2008;2:213-22.

32. Egerer K, Feist E, Burmester GR. The serological diagnosis of rheumatoid arthritis: antibodies to citrullinated antigens. Deutsch Arzteblatt Int 2009;106:159-63.

33. Slodkowska EA, Ross JS. MammaPrint 70-gene signature: another milestone in personalized medical care for breast cancer patients. Expert Rev Molec Diagn 2009;9:417-22.

34. Keedy VL, Temin S, Somerfield MR et al. American Society of Clinical Oncology provisional clinical opinion: epidermal growth factor receptor (EGFR) mutation testing for patients with advanced non-small-cell lung cancer considering first-line EGFR tyrosine kinase inhibitor therapy. J Clin Oncol 2011;29:2121-7.

35. Hughes SE, Barnes RO, Watson PH. Biospecimen use in cancer research over two decades. Biopreserv Biobank 2010;8:89-97.

36. Rudloff U, Bhanot U, Gerald W et al. Biobanking of human pancreas cancer tissue: impact of ex-vivo procurement times on RNA quality. Ann Surg Oncol 2010;17:2229-36.

37. Copois V, Bibeau F, Bascoul-Mollevi C et al. Impact of RNA degradation on gene expression profiles: assessment of different methods to reliably determine RNA quality. J Biotechnol 2007;127:549-59.

38. Saxena S, Sankhla B, Sundaragiri KS, Bhargava A. A Review of Salivary Biomarker: A Tool for Early Oral Cancer Diagnosis. Adv Biomed Res 2017;6:90.

39. Arellano-Garcia ME, Hu S, Wang J et al. Multiplexed immunobead-based assay for detection of oral cancer protein biomarkers in saliva. Oral Dis 2008;14:705-12.

40. Irimie AI, Braicu C, Sonea L et al. A Looking-Glass of Non-coding RNAs in oral cancer. Int J Molec Sci 2017;18(12). pii:E2620. Doi:10.3390/ijms18122620.

41. Ling H, Pickard K, Ivan C et al. The clinical and biological significance of MIR-224 expression in colorectal cancer metastasis. Gut 2016;65:977-89.

42. Li Y, St John MA, Zhou X et al. Salivary transcriptome diagnostics for oral cancer detection. Clin Cancer Res 2004;10:8442-50.

43. Hu S, Arellano M, Boontheung P et al. Salivary proteomics for oral cancer biomarker discovery. Clin Cancer Res 2008;14:6246-52.

44. Jou YJ, Lin CD, Lai CH et al. Proteomic identification of salivary transferrin as a biomarker for early detection of oral cancer. Analyt Chimica Acta 2010;681:41-8.

45. Brinkmann O, Kastratovic DA, Dimitrijevic MV et al. Oral squamous cell carcinoma detection by salivary biomarkers in a Serbian population. Oral Oncol 2011;47:51-5.

46. Kapeu AS, Luostarinen T, Jellum E et al. Is smoking an independent risk factor for invasive cervical cancer? A nested case-control study within Nordic biobanks. Am J Epidemiol 2009;169:480-8.

47. Poppe WA, Ide PS, Drijkoningen MP, Lauweryns JM, Van Assche FA. Tobacco smoking impairs the local immunosurveillance in the uterine cervix. An immunohistochemical study. Gynecol Obstet Investig 1995;39:34-8.

48. Suk K, Chang I, Kim YH et al. Interferon gamma (IFNgamma) and tumor necrosis factor alpha synergism in ME-180 cervical cancer cell apoptosis and necrosis. IFNgamma inhibits cytoprotective NF-kappa B through STAT1/IRF-1 pathways. J Biol Chem 2001;276:13153-59.

49. Yakob M, Fuentes L, Wang MB, Abemayor E, Wong DTW. Salivary biomarkers for detection of oral squamous cell carcinoma – current state and recent advances. Curr Oral Health Rep 2014;1:133-41.

50. Gulcher JR, Kristjansson K, Gudbjartsson H, Stefansson K. Protection of privacy by third-party encryption in genetic research in Iceland. Eur J Hum Genetics 2000;8:739-42.

51. Schmidt MK, Vermeulen E, Tollenaar RA, Van't Veer LJ, van Leeuwen FE. Regulatory aspects of genetic research with residual human tissue: effective and efficient data coding. Eur J Cancer 2009;45:2376-82.

52. Nietfeld JJ, Sugarman J, Litton JE. The Bio-PIN: a concept to improve biobanking. Nat Rev Cancer 2011;11:303-8.

53. Saha K, Hurlbut JB. Research ethics: Treat donors as partners in biobank research. Nature 2011;478:312-3.

54. Master Z, Claudio JO, Rachul C, Wang JC, Minden MD, Caulfield T. Cancer patient perceptions on the ethical and legal issues related to biobanking. BMC Med Genomics 2013;6:8.