Biobanking 3.0: Evidence based and customer focused biobanking

Daniel Simeon-Dubach a,⁎, Peter Watson b,c

⁎ medservice—Biobanking consulting & services, Hoerndlirain 22, 6318 Walchwil, Switzerland
b BC Cancer Agency, Victoria, BC, Canada
c Office of Biobank Education and Research, Dept. of Pathology, University of British Columbia, Vancouver, BC, Canada

ABSTRACT

Biobanking is a new and very dynamic field. To achieve long term financial sustainability of biobank infrastructures we propose that a new focus is needed on activities, products and services provided by the biobank that relate to the external stakeholder: biobanking 3.0. Earlier stages of biobanking are biobanking 1.0 (primary focus on the number of biospecimens and data) and biobanking 2.0 (primary focus on the quality of biospecimens and data). Both stages 1.0 and 2.0 are predominantly product oriented areas and have required a mostly internal focus on operational development within the biobank itself.

In this paper we will introduce our concept of biobanking 3.0 which capitalizes on the earlier stages but dictates a shift in focus to enhancing the value and impact for the three major sets of external stakeholders (people/patients, funders, and research customers) and creating a path to balanced and planned investment in biobank infrastructure and the sustainability of biobanking.

Biobanking 3.0 will improve real understanding as well as perceptions of value across different stakeholders. Patients and donors will appreciate seeing how their biospecimens and data are effectively used for research. Funders will value the ability to plan efficient targeting of funding and to monitor the impact of their support. Researchers will capitalize on the ability to translate their ideas into effective knowledge. Ultimately adoption of biobanking 3.0 will impact on the sustainability in the three main dimensions relevant to biobanking: social sustainability (acceptability), operational sustainability (efficiency), and financial sustainability (accomplishment).

Introduction

In a special issue of Time magazine in 2009 biobanking was featured as one of “10 ideas changing the world right now” [1]. Human biobanking has certainly evolved into a dynamic and complex activity. However a new period of metamorphosis is needed to allow the idea to emerge into a fully fledged discipline that is capable of continuing to drive research and discovery. The need to evolve further is in part attributable to persistent fragmentation and lack of harmonization around standards. But it is also needed because of the pressure of rising costs involved in current approaches to biobanking that threatens to outstrip the costs of previous forms of research infrastructure such as animal care facilities and optical imaging units [2] as demand continues to rise. In view of this there is an urgent need to consider a transformation of our approach to biobanking. Biobanking has thus far already seen at least two recognizable stages of evolution. But we propose that a new focus is now needed on activities, products and services provided by biobanks that relate to the needs of external stakeholders in order for the biobank infrastructure to become sustainable. In this paper we clarify our definition of a human research biobank and why it is worth investing in biobanking, then introduce the concepts advocating for a renewed focus: biobanking 3.0, and offer some solutions for implementation of the essential features of biobanking 3.0 to meet the challenges ahead in modern biobanking (Fig. 1).

It can be argued that modern biobanking has developed only in the last few decades. Consequently international biobanking organizations like ISBER [3], P3G [4], ESBB [5] and national and supranational structures like BBRB (Biorepositories and Biospecimen Research Branch of NCI, previously known as OBBR) [6], the Canadian Tumor Repository Network CTRNet [7] and the European biobank network BBMRI [8] were established in the early 2000s. It has also been just over a decade since international biobanking best practice guidelines were developed and published for the first time by some of the above-mentioned institutions and by the OECD, respectively [4,6,9–11]. Evidence of the dynamic nature and growth of biobanking has been the need for updates of best practice guidelines in a relatively short period. These activities have been paralleled by the introduction of new laws in many jurisdictions dealing specifically with biobanking.

The founding of these societies and the publishing of guidelines, best practices and recommendations, clearly indicate that the area of biobanking has become established as an independent scientific and professional entity. Professionals with different skills are also now working in biobanks. This includes among others those who can handle the large amount of data often involved, those in charge of all aspects of clinical biobanking of biospecimens, ethical and legal experts and last
but not least managers of biobank infrastructures. In parallel a new research area encompassing efforts to understand the influence of so-called ‘preanalytical’ variables on the research analysis of biospecimens has become attractive and important to explore. These variables include a range of possible patient, procedure, and tissue origin variables as well as collection, handling, and preservation variables that surround the process of biobanking biospecimens.

**Biobanks and biobanking**

There is as yet no general consensus about the definition and classification of the many different types of biological collections [12,13]. Several terms have been used interchangeably in the past [14] and a new classification based on function may now be more appropriate to categorize the diverse types, purposes and operational designs of collections [15]. The simplest term ‘biobank’ is widely used and in a recent survey of over 300 individuals across several international communities that identify themselves with biological collections of all types, a very broad definition of the term ‘biobank’ emerged [16]. It was proposed that “A biobank is a facility for the collection, preservation, processing, storage and supply of biological biospecimens and associated data, which follows standardized operating procedures and provides material for scientific and clinical use.” We consider here the subset of biobanks that collect human biological biospecimens and person related data to support human health research. Even within this subset, these biobanks are still very diverse. This is because biobanks are the physical hub of a complex set of activities that typically comprise the process of biobanking and the entire process is related to diverse lines of enquiry.
in health research (as depicted in Fig. 2). Biobanking typically starts when patients are asked to consent to donate their biospecimens and data for research. In some circumstances this step does not occur because a waiver of consent is provided by an ethics review board. The next step is to collect the biospecimens and data and to transfer these materials to the biobank. Within the biobank the biospecimens are accessioned, processed and either used directly in research or stored to preserve their integrity and quality as appropriate for an intended future research use. Data collection and storage within a database occurs in parallel to annotate the biospecimen at the time of collection and usually continues over time as an intermittent but ongoing process. All these processes are accompanied by governance mechanisms including most importantly mechanisms to guarantee protection of illegitimate access to personal data and to ensure utilization of the biospecimens and data when needed. The effects of variability in all these processes on the final use and value of the research product is now widely recognized and as described above is the focus of the new research area called biospecimen science that has emerged. So attention to careful replication of the processes and their documentation has become integral to the operations of high quality biobanks.

The majority of biospecimens stored in disease oriented tissue biobanks may be from cancer patients. In the medical pathway leading to a diagnosis of cancer, a biopsy or surgery is typically indicated and surplus tissue arising from these procedures can be stored in biobanks. These biospecimens have been proven to be very valuable especially for biomarker research and many of the best current examples of the value of personalized medicine based on guiding biomarkers, are in the field of oncology. At the same time, this implies that the impact of failure to transform the biobanking infrastructure, so that it can adequately support future discovery and validation of biomarkers, will limit clinical adoption of new drugs and therapies in this field [17,18]. To accommodate this diversity of functions biobanks range significantly in their design. For example biobanks may vary from simple to very complex structures, they may exist for a short time if biospecimens and their products are all used immediately after collection or alternatively are maintained for many years, and they may be standalone facilities or integrated into a research or a health care providing structure, respectively. The governance and funding can be both academic and commercial, and linked to different types of research organizations. Depending on the facility, some of the procurement processes may take place outside of the biobank. However as noted above, it has increasingly been recognized that documentation around controlled and standardized processes is critical for interpretation of research seeking to understand the biological differences between biospecimens and their original hosts.

A common biobank model therefore involves a capability for complex processing and long term storage of biospecimens and accompanying annotation with clinical treatment and outcomes of the original host. This requires a significant up-front capital investment in different tools and equipment that is needed as well as sustained management of the collection [19]. Therefore biobanks can become very valuable over time and this has recently focused special attention on disaster recovery plans to protect these investment [20–24].

**Focus on quantity: Biobanking 1.0**

In the first period of modern biobanking the major focus was on the number of biospecimens accrued and held “in stock”; we refer to this period as biobanking 1.0 (Fig. 1). This was the time when many biobanks were established by researchers who could not find the biospecimens they needed to progress with their own research. Expectations were that large numbers of biospecimens would be needed, in the thousands and tens of thousands, for many studies. For some the assumptions were that these needs were best met by private collections accessible only through collaborations. But it was also at this time when leading biobanks started to establish networks. The main driver being to reach a higher number of biospecimens in a shorter period to meet the increasing scale and complexity of selection criteria of research studies and to expand access. Those biobanks associated with genome-wide association (GWAS) studies are one example of biobanks developing within this era. Not surprisingly an analysis of the numbers of GWAS studies that can be retrieved using PubMed shows a dramatic increase in publications from around 100 papers per year in the first half of the last decade to around 3000 papers/year in 2010 (that appears to have leveled off since then). In fact it was one of the primary goals of the recently established BBMRI network in Europe to create a catalog that could allow an easy overview of available biobank resources such as these, along with the conditions under which they could be accessed, and contacts for access and the establishment of research collaborations [25]. Biobanks associated with GWAS type studies may however not be typical and the majority of biobanks and studies with biobanking components have needed much smaller numbers of biospecimens for their research projects. But there has also been significant growth in the biospecimen cohort sizes used irrespective of biobank size. As shown in a recent study [26], the average number of biospecimens used in cancer research studies increased from about 50 biospecimens per paper to between 150 and 200 in recent years. Similar figures have emerged from our own analysis of a set of requests from researchers working in the biotech industry.

So how large is a typical biobank? Two newly published surveys of biobanks in Europe and the US showed similar results [27,28]. About one quarter of biobanks were small (defined in these studies by a cut point of <1000 biospecimens), 30% had 1000 to 10,000 biospecimens, 22% and 13%, stored between 10,000–50,000 and 50,000–100,000 biospecimens, respectively and the rest were larger biobanks or did not report the number. Another published estimate concluded that at a minimum there are 8000 small, 2000 medium, and 80 large scale biobanks (defined small/moderate/large using cut points of <200, 200–1000, and >1000 cases, respectively) across Canadian medical centers alone [15]. The small sizes of many biobanks may come as a surprise. However many typical disease oriented biobanks begin with a specific hypothesis that stimulates an effort to compile a prospectively collected biospecimen cohort. These are also usually funded within the context of basic discovery or translational research studies. For these, often small to medium sized biobanks, pragmatic funding limitations, urgency to generate publishable results, or statistical power calculations determining optimal study sizes are among the several reasons why collections are usually associated with a planned target size. Similarly some large biobanks (for example large, population based blood biobanks designed to support research into disease risk factors) begin with specific or general hypotheses and careful planning around target numbers, partly stimulated by the prospect of very large sizes and need to plan for storing hundreds of thousands of biospecimens. But more often small biobanks evolve to operate without a strategic plan. In the European survey mentioned above 57% of biobanks reported that they have no specific sample target numbers [27]. These biobanks arise out of a desire to prospectively compile collections that can support multiple forms of “retrospective hypothesis research”. That is they serve the function of enabling researchers to address questions that involve selecting biospecimens associated with historical annotating data (e.g. patient treatment and outcome data). While good arguments exist for supporting some biobanks to operate in this way, it is important for the research and biobank community to rationalize the numbers and scale required for this type of biobank model.

In 2000, the Rand Corporation estimated that more than 300 million human biospecimens (e.g., formalin-fixed paraffin-embedded tissues, frozen tissue and fluids) had been stored across the United States, with this volume growing at a rate of about 7%, or approximately 20 million specimens, annually [29]. However this growth in numbers of stored biospecimens should be considered alongside recent opinions expressed by biobank managers in the literature [30] and our personal experience that typically only between 10% and 50% of biospecimens stored in tumor tissue biobanks will ever be used.
Focus on quality: Biobanking 2.0

However as quantities of biospecimens available for research have expanded many crucial and interrelated issues have emerged and are facing the existing biobank infrastructure including uneven quality, lack of evidence-based standards and adequate tools to enable their implementation challenges to sustainability, and often lack of awareness among biomarker researchers accessing biobanks and their stocks, of the specific quality requirements for biomarker development and registration process. In this second stage of modern biobanking a major focus on the quality of biospecimens accrued has emerged; we refer to this stage as biobanking 2.0 (Fig. 1). The drivers for this new focus are several as discussed in the following sections.

Reproducibility of research based on biobanks

Many researchers now question the validity of their past findings because of concerns about the quality of the biospecimens [32]. While this issue has been known for a long time [33] to biobankers, recent papers have highlighted this issue for the broader scientific community as it relates to the very low reproducibility of scientific data published even in high impact factor journals [34,35].

Reasons highlighted for this lack of reproducibility [34] include several aspects of statistical design and analysis, and insufficient description of materials and methods. Other reasons include poor study design (such as insufficient sample size), bias towards publishing positive results, and pressure and competition among laboratories to publish. Although the visibility and prestige of journals in which a study is published positively influence the citation rate, inadequate statistical reporting quality and inappropriate statistical analysis do not appear to affect dissemination of published medical science [36]. Another issue cited is the inadequacy of the peer review system in general. This system relies on reviewers who have limited time and no resources to reproduce data: so errors often remain undetected [37].

Also many papers where good peer review has identified design and statistical issues may be rejected initially but then will subsequently be published in other journals without substantial changes or improvements [38]. Even more worrisome was how many times the primary papers have been referenced without validation [35]. From the 53 papers that were analyzed in the study by Begley et al. 21 were published in journals with an impact factor >20. Six papers could be reproduced and were evenly distributed among the two groups defined by the journal impact factor (personal communication with [35]). The number of citations of 18 articles published in high impact journals and defined as non-reproduced (e.g. not being sufficiently robust to drive a drug development program) ranged between 3 and 800. The citations for the other 29 non-reproduced articles published in journals with impact factors from 5 to 19, ranged between 6 and 1909. The author proposed several steps including trying to change the general attitude not to publish negative data. Funding agencies, reviewers and journal editors must agree that negative data can be just as informative as positive data. Journal editors must play an active part in initiating a cultural change. Researchers should be able to discuss and report troubling or unethical behaviors without fearing adverse consequences. When analyzing the scientific literature about clinical biobanking of biospecimens (e.g. data about transport, receiving, storage, annotation, processing, quality control, and release/distribution of biospecimens) (Fig. 2) we identified 125 papers in PubMed by using the keyword “biomarker discovery” that were published in 2004 (31 papers) and 2009 (94 papers) [39]. One would expect papers on biomarker discovery to clearly describe the biospecimens that were used. In more than 50% of cases there was no information whatsoever about the biospecimens used for the research project to reliably validate the findings. When approaching the editors of 60 biomedical journals that focus on relevant topics such as biomarkers, biobanking, cancer, pathology and laboratory medicine, none were interested in further investigation into how to improve the reporting of the information about pre-analytical variables associated with biospecimens [40]. Still a reporting standard called Biospecimen Reporting for Improved Study Quality (BRISQ) on relevant data to include when publishing papers already exists [41] and this standard has been implemented by some journals [42–44]. This standard will encourage biobankers to invest more in collecting data about pre-analytical variables in the future. Otherwise they will risk losing their value to future researchers, who may specifically search for biobanks that will be able to supply this kind of biospecimen related data in order to be assured of being able to publish their results. We also expect that if the information required by the BRISQ guidelines is to be included in published papers that the percentage of papers with reproducible findings could increase. If we are not able to solve the issue of irreproducibility then some experts fear that the scientific literature will continue to be polluted with incorrect data and that it will take a long time to sort out and significantly blunt “the war on cancer” [45]. This could have a serious impact on the credibility of all health research. Researchers have of course already learnt to adapt to some of the limitations of current biobanking and difficulty obtaining biospecimens that meet their needs [32]. Restricting the scope of their work to cell lines and other models of human disease, and then translating these findings to small often biased sets of biospecimens, at least mean that it is possible to arrive at conclusions that can be published. To overcome this crisis researchers need to be able to access biospecimens and data of defined and adequate quality, and to be able to select materials from biobanking infrastructure that is representative of populations if we
are to realize the benefits of much of the current investment in health research. Hence investment in biobanks and solutions to biobanking are justified. There is another aspect. The potential for not being able to reproduce academic data is a disincentive to early stage investors. Venture capital firms have now started hiring CROs to independently validate academic science prior to putting up serious money [46].

Economics of biobanking

The importance of biobanking is recognized within many spheres including social health research, economics, and the public [1,47,48]. As just one example the US National Cancer Institute spends over $50 million/year on biobanking [47]. Nevertheless there has been relatively little attention given in the literature to the economics of the activity, and where this has been considered different approaches have been used. One is to calculate the cost of biobanking and different cost recovery models. The NCI in the US conducted and published a series of papers on this topic, referred to as biobankonomics [2,19,48]. Vaught et al. estimated the first year start-up costs for a biorepository of 50,000 biospecimens to be from $3 million to $5 million [2].

A biobanking for academic and industrial research and development [31]. Using an input–output model that computes regional economic multipliers, it was estimated that a mid-sized biobank could create a yearly economic impact of $20 million (personal communication).

A report by the UK based STRATUM (Strategic Tissue Repository Alliances Through Unified Methods) stressed the importance of biobanking for academic and industrial research and development [31]. However financing and organization of biobanks in the UK are believed to no longer meet the needs of the scientific community. The survey they conducted showed not surprisingly a very diverse landscape of biobanks comparable to other countries and regions. It also became clear that biobankers in general are not fully aware of the costs of biobanking. According to STRATUM it is neither possible nor desirable to propose a standard cost model for biobanking mainly due to this diversity of biobanks. The report concludes in favor of public funding with additional private funding being ideal. The host (public) institution of a biobank should provide the majority of funding enabling investments in best practices and guarantees of financial sustainability, and users should only pay a marginal part of the cost for actual release of biospecimens and data. Although this would be ideal and worth working towards, we doubt whether this approach is feasible in the short term, bearing in mind the economic situation in many European and North American regions. Systematic deployment of public money for an organized biobank infrastructure is certainly desirable but as just one of several types of research infrastructure this requires a discussion on the balance to be struck between biobanks and funding allocations and scale of other existing infrastructures and the research agenda itself. On a scientific level large biobanks and cohorts can currently be criticized for lack of explicit data on their impacts and lack of strategic plans for their targets and resource requirements that could grow to limit other types of research [50]. The general strategy of some large and costly biobanks has certainly been questioned [51].

According to two different reports the world market for biobanking in human medicine is predicted to be worth $23.9 billion by 2015 [52] and $24.4 billion by 2017 [53], respectively. The major stakeholders for these biobanking stocks and services are the pharmaceutical and biotechnology industry and academic research as they depend more and more on human tissue research for example in their preclinical drug development and discovery research. There is no doubt then that these stakeholders will need to consider the fact that biobanking will continue at least in the short term, to become more expensive as the desire and requirements to invest in better quality will need significant investment.

Complexity of data and biospecimen quality requirements

There is also no doubt that across biobanking there is value in an economy of scale and that to run a biobank successfully a critical mass of biospecimens and health data should be collected. It is not efficient to have every researcher running his own collection in the way that it perhaps used to be in many institutions [54]. Although there will always be specific merits and a need to accommodate some relatively independent biobanks within an institution (e.g. to address special governance or stakeholder requirements), one centralized biobank, often functioning as host for several biobank units and/or the supporting hub to independent but linked specialized biobanks has increasingly proven to be a favored model adopted by institutions to serve their research communities. This allows hiring, cross training, and communication and technical interactions of fully dedicated staff who can manage all different aspects of modern biobanking. Also the high investment in tools and equipment that is mandatory for attaining the desired quality levels of biospecimens and data can be better justified and amortized across larger integrated collections. In the business plan of the biobank it should be clearly stated what type of data should be collected. As depicted in Fig. 3 there are four different types of data. These encompass person related and socio-economic data from donors and patients. Some of this data is strictly mandatory like gender, age and other very basic information. It is also helpful if links are known to enable how and where to get additional or validating data if needed. The types of data also encompass medical data including specific biological data in connection to the stored biospecimen(s) but also all other medical conditions. As with the person related and socio-economic data it is not important that all the data is stored at the outset in the biobank database but it is desirable to have a minimum data set to facilitate initial case selection for studies and to have a link to get further data quickly if needed. This additional search for both types of patient/donor related data is often done by the biobank as an additional service and for a predefined fee to enable the biobank to access secondary linked data resources. However this operational model is still often much cheaper than accurately compiling large data sets within the biobank database that may never be used or that need to be
updated at the time of selection of individual specimens for research. The other two types of data are related to the biospecimens. In the material transfer agreement it should be stated whether and if yes what kind of data generated when conducting a research project with biospecimens should be fed back to the biobank. Repatriating and accurately curating raw research data from every study accessing a large biobank is beyond the resources and capacities of most biobanks. But at a minimum a rough outline of the research should be added to the annotation of the biospecimen including information about the researchers and overview information about the project. This allows the research findings connected with each biospecimen to be located and potentially used by another researcher at a later stage. The final type of data that should be collected with the biobank is the data about the pre-analytical variables of the biospecimens, the quality control data and all other data generated when in-house biospecimens research was performed.

Pathologists have long known about the heterogeneity of primary cancers and the evolution that commonly occurs in secondary tumors arising over time with treatment and progression. Insights into tumor biology, such as intra-tumor heterogeneity, tumor–host crosstalk, and the evolution of the disease during therapy, require access to biospecimens from the primary tumor and also those that reflect the patient’s disease in specific contexts. This intra-tumor heterogeneity recognized between and within primary tumors includes both epigenetic and micro-environmental factors that also affect the cancer cell heterogeneity [55]. While early detection or prevention remains the goal in many areas of cancer research, the challenge in many cancers lies in treatment of metastatic or recurrent disease. And yet the majority of biospecimens in existing tumor biobanks are derived from surgical samples of primary tumors [56] and vanishingly few are representative of recurrences and metastases and later stages of disease. This has led to an important gap in our knowledge of tumors. For example, multiple papers have analyzed many aspects of the biomarker status of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor type 2 (HER2) receptor [57–62] in primary breast cancer. Metastatic breast cancers have historically been presumed to have the same predictive biomarkers as the initial primary tumor [59] and for the most part treatment decisions based on the biomarker status of the primary tumor are effective. But as shown by a large prospective study there are frequently discrepancies in ER, PR or HER2 status between primary and metastatic stages which when known can change therapeutic management for one in six patients [58]. One of the publications also reviewed the literature and found 32 relevant studies that investigated the concordance of the hormone receptors (ER and/or PR) and HER2 after neo-adjuvant chemotherapy with or without Trastuzumab [60]. Studies that concluded that ER and/or PR remained stable after neo-adjuvant chemotherapy were performed with evidently lower number of patients compared to studies that reported a change. A switch to a negative HER2 receptor status in up to 43% of the patients was reported when neo-adjuvant chemotherapy was combined with Trastuzumab.

The challenge of collection of biospecimens from many tissues and over time from the same patient raises a number of significant issues including ethical, medical, and organizational issues that have encouraged tumor biobanks to focus mostly on large primary tumor specimens. But the results of all the above-mentioned studies clearly show how important it is for research and therefore the biobanks that serve and enable research to collect not only one biospecimen of a primary tumor but also a collection of biospecimens from different locations and at different time points wherever possible.

**Biobank value and disaster recovery**

Although not a driver of the focus on quality in biobanking 2.0 an important outcome is the significant increases in costs as discussed above, and therefore value. A common biobank model now involves a capability for complex processing and long term storage of biospecimens and accompanying annotation with clinical treatment and outcomes of the original host. This increasingly requires a significant capital investment in different tools and equipment that is needed for initial establishment as well as sustained management of the collection [19]. And the figures above clearly indicate that the vendors of these different tools and equipment also anticipate that biobanks will invest accordingly. Therefore biobanks can become very valuable over time and this has recently focused special attention in the biobanking 2.0 stage on disaster recovery plans to protect these investments. Over the past several years major natural (meteorological, geological, or biological), human (accidental or intentional), and technological (power, telecommunications, hardware, and software) disasters have been reported affecting biobanks at a higher prevalence than ever before. This increase may be due to better accounting of incidents but also appears to be due to an escalation in
frequency of significant natural catastrophes. Awareness of the potential for risk in biobanking and creating contingency plans is a core competency all biobanks must adopt. The responsible question is not ‘will the biobank be hit by a disaster?’ but ‘will the biobank be ready when a disaster occurs?’ These disaster recovery plans have to follow location specific conditions. An adequate risk assessment and action & contingency plan need to be established to prevent or at least mitigate the consequences of natural and human disasters like fire, flooding, earthquakes and hurricanes [20–24].

Focus on evidence and on customers: Biobanking 3.0

In considering the stages of biobanking it is important to acknowledge that human research biobanks have existed in primitive forms, perhaps best described as ‘organized collections’, for more than 50 years. These collections were suitable for personal or very local use in an era before ‘electronic communication’ and ‘high throughput technologies’. In the late 1980s the first generation of recognizable biobanks emerged in response to the simple demands for increased numbers of biospecimens and improved oversight. At the same time an exponential growth of newly discovered potential biomarkers was noted that also had a significant impact on the increased demand for biospecimens and data. Further desire for improved biospecimen quality through standardized processes and annotation led to the second generation of biobanks that have been developed in relatively small numbers since the early 2000s [28]. The operational focus of each of these generations of biobanks can be categorized as discussed in sections above as biobanking 1.0 (primary focus on the number of biospecimens and data) or as biobanking 2.0 (primary focus on the quality of biospecimens and data). These areas of primary focus in stages 1.0 and 2.0 are predominantly product oriented areas and have required a mostly internal focus on operational development within the biobank itself. Biobanking 3.0 not only capitalizes on these earlier stages but also dictates a shift in focus to enhancing the value and impact for the three major sets of external stakeholders (people/patients, funders, and research customers) and creating a path to balanced and planned investment in biobank infrastructure and the sustainability of biobanking (Fig. 1).

One new element of biobanking 3.0 is implementation of mechanisms to enhance the enrollment of donors (people or patients) and broader access to their biospecimens. An example that allows direct engagement of all patients attending health facilities is a ‘Permission to Contact’ platform to enable the consent process for biobanks [63]. Another approach is to directly involve the patients in the consenting process for biobanking [64]. These mechanisms allow biobanks to be efficient and approach the most relevant patients such as those who undergo medical procedures that yield biospecimens. This not only can increase potential enrollment rates but also decrease the costs and so impact biobank efficiency [63]. It also strengthens the knowledge of the public and acceptability around biobanks.

Another element of particular interest to funders is development of a plan. Few biobanks in the previous phases of biobanking have been started with a biobank strategy based on market research and surveys or specified targets or consideration of data on trends in biospecimen use, a business plan, or plans to obtain customer satisfaction information. And yet this is vital information and is the basis that can allow a biobank to focus its resources and to develop additional services, e.g. advise in trial design in general, council researchers on the best biospecimens to be used for the specific study and make sure all requirements for biomarker R&D are followed correctly, minimizing issues arising from technical failures when submitting the biomarker dossier for registration. Consideration of the well documented example of the disappointment faced early on by “The Cancer Genome Atlas” (TCGA) project is instructive when the biospecimens held by existing biobanks were found to be ‘inadequate’ [65]. However leaving aside questions around what level of funding had been available to support the biobanks that were approached or the practicality of the selection criteria to determine ‘inadequate’, the episode illustrates that many biospecimens have been collected without consideration of the needs of the customer and may therefore never be used. All biobanks should therefore have a detailed strategic plan that should also include how to market the infrastructure and the products offered by the biobank and encourage research use. Examples of services some biobanks already provide to address this are advice on research design and which biospecimens and data will fit best for a specific research project, lease of part of the storage capacity to outside research groups or institutions, and emergency backup support for other biobanks.

An important element related to developing and monitoring a strategic plan is implementing common internal and external performance metrics to inform decisions by both the biobanks themselves and funders of the infrastructure. Since in many cases the biobank funders are also funders of the research end users, the extent of support should be a natural balance. The reason it is not may be largely because of the difficulty in measuring biobanks. A successful biobank needs to consider an appropriate balance between elements such as an existing stock of biospecimens and mature patient outcome data, the proportion of non-renewable biospecimens kept for future use, its capability to replenish this stock, and the outflow of biospecimens transferred to researchers. Outflow partly depends on building a stock that is of value and this requires some planning (e.g. around scale of biospecimen collections that will be adequate to accommodate the application of typical selection criteria in different disease areas) and some educated guesswork. Measures of biobank value should therefore relate more to the outflow and the impact rather than the size and stock [66]. Managing the biospecimens and data on stock will also optimize the workload and cost of every step in the clinical biobanking process. One approach to metrics is to implement certification and accreditation processes that provide conduits to educate biobanks on common processes including metrics [67] and external assurance. Both should also be relevant to funders.

The research customer comprises the third group of external major stakeholders. Researchers’ current requirements should define focus and extent of accrual, processing and annotation, and their future needs and demands the scale of storage. Some quantitative and qualitative data in these areas is available. Examples are the trends in the average number of biospecimens of different formats needed for publications or the applications for these biospecimens and their products [26,28,68]. It should become standard and part of every material transfer agreement for all research users to assume the responsibility of reporting back the results produced (at least in general terms including quality assessments, amalgamated data, and publication outputs), while using biospecimens and data received from a biobank. The sharing between biobanks of experience with repeated contacts with researchers is one possibility to develop an in-depth knowledge about the biobanking market and the requirements of its different research stakeholders. But requirements are partly driven by the specific questions and available levels of funding that dictate scale of cohorts that can be assessed, but also by pragmatic factors such as requirements set by journals. The development and gradual increasing adoption of guidelines on data that should accompany research papers describing findings based on patient cohorts (e.g. REMARK) [69] and their biospecimens (e.g. BRISQ) [41], creates a new driver for what data annotation is important to researchers [42,44]. Proof of the professionalism of the biobank through accreditation and certification processes may also become a required component in higher impact journals. At the same time, if biobanks can respond by collating and providing these kinds of data to researchers in an efficient format then it becomes a new product valued by its customers. As mentioned above guiding researchers to prevent the pitfalls in the process of biomarker registration will make a biobank more attractive to researchers from both academia and industry. All these activities will eventually strengthen the professional relationship between the biobanker and the different researchers.
In conclusion, a new focus on biobanking 3.0 will improve real understanding as well as perceptions of value across different stakeholders. Patients and donors will appreciate seeing how their biospecimens and data are effectively used for research. Funders will value the ability to plan efficient targeting of funding and to monitor the impact of their support. Researchers will capitalize on the ability to translate their ideas into effective knowledge. Ultimately adoption of biobanking 3.0 will impact on the sustainability in the three main dimensions relevant to biobanking: social sustainability (acceptability), operational sustainability (efficiency), and financial sustainability (accomplishment) (Fig. 4).

References

[22] OECD. OECD guidelines on human biobanks and genetic research databases; 2009.


Patient public involvement (PPI) in consenting for biobanks. Biopreserv Biobank 2013;11(3) [A-4 ELSI 06].


Certification for biobanks: the program developed by the Canadian Tumour Repository Network (CTRNet). Biopreserv Biobank 2012;10(5):426–32.
