

A Rational Framework for Development Iranian Twin Registry: Lessons Learned from Developed Twin Registries

Hamidreza Abtahi

Tehran University of Medical Sciences

Marsa Gholamzadeh (✉ m-gholamzadeh@razi.tums.ac.ir)

Tehran University of Medical Sciences <https://orcid.org/0000-0001-6781-9342>

Leila Shahmoradi

Tehran University of Medical Sciences

Mamak Shariat

Tehran University of Medical Sciences

Kazem Zendehtdel

Tehran University of Medical Sciences

Research article

Keywords: Twin registry, Conceptual model, Workflow, Information flow

Posted Date: July 21st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-44309/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Regarding technology advancement in health sciences and the possibility of recording and organizing a huge amount of data, registries in various clinical domains have been established. The specific genetic structure of twins has enabled researchers to find answers to role of genetics and environment in medical sciences. Thus, twin registries were developed across the world to support twin studies. The main objective of this study is devising a conceptual model for developing the Iranian twin registry to ensure the success of this registry.

Methods: In this descriptive and qualitative study, the combination of literature review, brainstorming, and focus group discussions were applied to develop a conceptual model of the twin registry. Based on our qualitative and thematic analysis, the workflow of information gathering and different data levels are described in this study.

Results: According to the characteristics of established worldwide twin registries, the success factors in registry implementation were recognized. Moreover, based on our objectives, the hierarchical conceptual model of the registry was invented. The source of information, the different levels of data, and the information flow were determined based on this model. The workflow modeling of data gathering was represented through this survey to clarify the data collection and saving processes.

Conclusion: Suggesting a conceptual model and standard framework for twin registry development at the national level based on the experiences of other countries could guide researchers in twin registries implementation efficiently. Moreover, the next steps of implementing a twin registry in a standardized and stepwise way are outlined in this project.

Background:

Twin studies provide valuable means for clinical researches to study a wide range of health topics [1–4]. The specific genetic structure of twins enabled researchers to find the answer to some puzzling questions in medical sciences by studying the impact of genetic and environmental factors in different domains of medicine using the health-related data of twins and their long-term clinical follow-up [5, 6]. It is necessary to conduct a large clinical trial and longitudinal cohort studies to achieve such valuable results [7, 8]. With the advancement of healthcare technology and the possibility of using large data repositories, clinical registries have been invented to overcome the challenges of traditional clinical studies [9, 10].

Clinical registries provide the opportunity for researchers to aggregate health-related data from different resources into a large database to support intelligent data analysis and mass screening [11, 12]. They are known as an integrated system that gathers information using structured minimum data sets from different sources [13]. By equipping registries with knowledge discovery capabilities, researchers can extract valuable information and even generate new assumptions in different fields of medical science [2, 14]. Twin study is one of the disciplines in which creating registry has been so beneficial [15].

In the last three decades, the national twin registries have been founded in developed countries to create a national database containing primary and secondary data related to identical (monozygotic) and non-identical (dizygotic) twins [8, 16]. In addition to other benefits, organizing and recording information of twins would be decreased the bias of the population-based survey [4, 7, 17, 18].

Developing a twin registry at the national level required outreach efforts. Indeed, developing a twin registry encountered many challenges in the context of designing, gathering information, registration, follow-up, and implementation [15, 16, 19].

One of the most important challenges in registering twins is the difficulty of collecting twin's health data. Because the twin's information cannot be accessed through health information systems [19]. The next challenge is related to the complicated nature of developing any national registries. Morsey et al. [20] highlighted the challenging nature of developing health registries in their study. They proposed that the initial steps for developing the national registry are defining the proper strategic plan for the whole process and designing the conceptual model for information flow. Lack of attention to these issues led many countries to fail in their efforts to develop national registries. So, main objectives of this study are: (1) defining a strategic plan for twin registry implementation, (2) devising a conceptual model for developing the Iranian twin registry, and (3) generating a suitable workflow process to ensure the success of this registry.

Methods:

In this descriptive and qualitative study, the combination of literature review, brainstorming, and focus group discussions were undertaken to answer our research questions. The applied methodology can be described in four phases. These four steps are (1) literature review, (2) focus group (FG) meeting to gather expert opinions, (3) thematic analysis and synthesis of the results, (4) summarizing the results and generating the models.

Literature review:

In the first phase, the literature review was conducted based on the identification and reviewing of the challenges, pitfalls, and achievements of twin registries throughout the world. Scientific database searching was performed using PubMed, Scopus, Web of Science, BMJ, and ScienceDirect. The search strategy included keywords such as "Twin registry", "Twin registry studies", "National twin registry", "Twin registry outcomes", "Twin studies in the world", "Twin registries". The official websites of developed twin registries were also searched to identify their characteristics.

All founded registries were explored in terms of the country, type of study, primary goal, population, their method for information gathering, biobank samples, population, kind of registries, and outcomes. Our criteria were limited only to articles regarding the designing of twin registries and their characteristics.

Focus group meeting to gather expert opinions

Following the literature review, a FG technique was employed to discuss the foremost requirements for Iranian national twin registry implementation. The experts were chosen by purposive sampling method based on expertise and the objective of the study. Then, the researcher (MG) contacted them by phone and email and explained the objective of project, and asked for their participation. Out of 15 experts, ten experts agreed to participate in our FG discussions.

The date of the meetings specified according to the expert's wishes at the Tehran University of Medical Sciences. In these meetings, the participants were asked about their agreement on essential requirements and characteristics of the Iranian twin registry.

Based on the literature review results, different levels of information, predefined categories, our objectives of conducting this research was explained thoroughly at the beginning of the meetings. Then, the main requirements and qualitative analysis were discussed in the FG discussion meetings. The FG discussions allowed the investigators to gather expert opinions and discuss the research questions face to face. The FGs typically lasted 60–90 minutes, although the length of meetings varied according to topics and the participant's contribution during each session of the discussions. Before each discussions, the objective of our study was explained to participants, and verbal consent was obtained. One of the researchers (M.G) as a Ph.D. student who had experience in designing health information systems and medical researches took notes to capture the significant points of discussion. Transcripts were checked by other researchers at the end of sessions. Key points and quotes from each FG were obtained by transcription. FGs were continued until to reach data saturation. Thus, five FG discussions proved to be sufficient.

Thematic analysis and synthesis of the focus group results

Thematic analysis as a qualitative analysis method was applied to analyze the FG results and literature review findings. In this study, thematic analysis was conducted based on the method that was introduced by Braun and Clarke [21]. Atlas ti V.6 software used to manage all of the transcriptions. First, one of the researchers recheck transcriptions and FG data. Then, all of the FG data were imported into the software. Secondly, open codes assigned to notes with interpretive notes by two researchers independently. Third, all of the data were combined and the coding list was generated. Based on the repeated codes, searching for main themes was conducted to identify themes. Potential themes were reviewed by all authors to ensure their repetition. Finally, a thematic map was designed to generate an emergent theme and subtheme based on Braun and Clarke suggestion.

Summarizing the results and generating the models

In the third phase, all of the findings were categorized and summarized by a combination of literature review findings with expert panel results. Based on the main objectives of our study, the main

stakeholders, the information resource, and the workflow of information gathering were determined. All of the obtained information was classified and analyzed using SPSS. Following literature review and five FG discussions, the draft of the expected specifications was produced. Subsequently, the first version of the conceptual model of Iranian twin registry was produced and refined based on the expert's view.

After model development, two FG discussions were held to validate the reliability and accuracy of designed models. Additionally, the methodology and various steps of developing a twin registry were established through this phase.

Results:

Expert panel:

The multidisciplinary expert team included ten experts comprises of one internist, one pharmacotherapist, three epidemiologists, two pediatricians, two midwives, and one medical informatics specialist. All expert panel members participated in FG sessions. The demographic characteristics of all members represent in Table 1.

Table 1
The demographic characteristics of experts

	Data	Frequency	Percentage
Specialty	Internal medicine	1	10.0%
	Epidemiologists	3	30.0%
	Medical informatician	1	10.0%
	Pharmacotherapist	1	10.0%
	Pediatricians	2	20.0%
	Midwifery	2	20.0%
Age	30–40	2	0.2
	40–50	2	0.2
	50–60	6	0.6
Experience	< 5yrs	3	0.3
	5–10 yrs.	2	0.2
	> 10 yrs.	5	0.5

Characteristic of established twin registries based on literature review:

More than 2500 studies were retrieved in searching the scientific databases phase. After duplication removal, 2300 articles remained. In the title and abstract screening phase, approximately 530 articles remained. Finally, 39 studies were selected as relevant articles through full-text screening based on the main objective and research questions of our study. Additionally, all of the information about established twin registries was extracted from gray papers and their official websites. All of the retrieved studies were analyzed and classified based on predefined categories.

According to reviewed articles, 32 twin registries at the national or state level were determined across the world. The investigation disclosed that most of them were founded in developed countries. Due to the high potential of twin registries as the statistical and epidemiological research tool, the rate of constructing twin registries are increased in the last decades.

Out of 32 recognized registries, almost 76% of found registries are national twin registries, and 7% of them are provincial. The rest of the registries have belonged to non-governmental organizations. Regarding the type of study, 79% of them applied a longitudinal design. The oldest twin registry is the twin registry of Denmark (DTR) with more than 86 years of history[22]. Some of the twin registries are classified as volunteer registries. Australian Twin Registry (ATR) as the largest volunteer twin registry successfully registered 17% of identical and non- identical twins of Australia[23].

Today, 68% of twin registries equipped with Biobanks. They use different samples from twins based on their objective. The analysis of samples based on the type of sample shows in Table 2. All of the registries assessed the zygosity in different methods. Those techniques include employing a standard questionnaire, biological markers, or mixed methods of standard questionnaires and biological markers. Some of them applied decision tree rules in questionnaire-based methods. One of the most famous questionnaires used in twin studies is known as “peas in pod questionnaire”.

Table 2
The samples category of biobank

Sample	Used%	Numbers of registries Collect these samples
DNA	53.6	15
Plasma	17.9	5
Whole blood	17.9	5
Buccal cells	21.4	6
Serum	21.4	6
Salvia	28.6	8
Urine and stool sample	14.3	4

The survey showed that developed countries such as Sweden [24], the Netherlands [25], England [26], Finland [27], and Australia [23] founded their national twin registries to provide a reliable platform to enhance longitudinal twin studies. Concerning outcomes, the twin registry studies cover a wide range of clinical and non-clinical disciplines. They reached novel findings in various health topics such as finding new methods in the diagnosis of cancer, discovering causes of premature deaths, studying on fetal abnormalities, and determining underlying hereditary factors in chronic diseases, the efficacy of new drugs, and identifying risk factors of behavioral disorders [3, 8, 15, 28–30]. The development of a twin registry in some countries also led to the publication of more than fifty research articles annually. All the results obtained from the analysis of the characteristics of the developed registers were summarized in a table by authors. Then, the results were visualized as a network of the most frequently used words in Fig. 1.

The establishment of twin registries and conducting more studies in developing countries are needed for identifying specific genetic and environmental variables that contribute to health and disease. Thus, developing the Iranian twin registry would be highly valuable and beneficial in Asia.

The analysis of the reviewed studies revealed that the most important issues in the field of development and implementation of twin registers are the standardization of data types, information flow, data collection methods, storage methods, data analysis, required reports, and data linkage. All of the studies emphasized that data gathering should be done based on the standard workflow to enhance accuracy and outcomes.

In terms of software or platform, most of the registries did not refer to the type and description of the platform employed to collect twins' data. Thus, the researchers contact some registries to check the existence of these platforms via email or phone. According to studies and available results, about 65% of registries used information management software to organize their information. Forty-five percent of institutions developed these systems by themselves in their institution, and the rest of the registries

ordered their systems to software companies as needed. None of the existing registries stated that they used downloadable applications or open-source software because they believed that they did not meet their requirements.

Thematic analysis and discussion findings:

Different views and opinions represented regarding the design and implementation of the twin registry at the national level in FG discussions. The participants pointed to obstacles, opportunities, requirements, and challenges to achieve successful implementation. All topics were organized by key themes by thematic analysis. When no new aspects emerged, data saturation occurred, and the last group discussion took place. A total of five main themes were extracted from the thematic analysis of the FG discussions. Table 3 shows these main themes along with quotations from each topic.

Table 3
Summary of themes and applicable quotes based on the focus groups

No	Theme	Quotation
1	Lack of a proper strategy plan	<p>Expert 1: "To achieve our main objectives, the standard strategic plan should be defined"</p> <p>Expert 2: "It is not possible to start a long-term study at the national level without a codified program"</p>
2	Various data sources	<p>Expert 1: "Because twin information must be gathered from a variety of sources, it must be determined before studying."</p> <p>Expert 2: "Different levels of information and data at the national level and how different levels of data communicate before data collection should be defined"</p>
3	Define classification for twins	<p>Expert 1: "Since twin information will be collected from different age and social groups across the country with different characteristics, proper information classification in different age groups can help to implement the registry more successfully."</p>
4	Information workflow	<p>Expert 1: "The data collection process should be defined clearly. So that a team of different professionals can accumulate data according to a standard plan."</p> <p>Expert 2: "If there is no specific plan, everyone can collect data in their own way and it might cause data redundancy."</p> <p>Expert 3: "Without a standard workflow, invalid data might be entered into the database. Therefore, a proper solution must be considered to prevent the entry of invalid data."</p>
5	Conceptual model	<p>Expert 1: "Since there is no specific conceptual model for twin registers in the world, devising a conceptual model can overcome the obstacles and future barriers to registry implementation, and it could ensure the success of this system."</p> <p>Expert 2: "Developing a conceptual model could be considered as a roadmap for future planning."</p>
6	Data dictionary	<p>Expert 1: "A set of information that should be recorded in the system is one of the important requirements of any registry."</p> <p>Expert 2: "According to the initial aims of the registry, the required data in twin researches should be determined and explained."</p> <p>Expert 3: "The required data must be approved by experts."</p>

Based on these extracted themes and lessons learned from founded registries, the steering committee planned to design some rational models for creating a population-based twin registry with accessibility to twin's health data. Consequently, the suggested models to saturate the requirements are described in the following. However, data dictionary development and determining minimum data sets not discussed in this article, and it needs further study.

Developing a twin registry strategic plan:

The examination of established twin registries throughout the world helped us to determine the key features of a twin registry. Alongside, the feasibility and the strategic plan of developing a twin registry in Iran were discussed by researchers. All of the processes that should be performed for twin registry development are represented in Fig. 2. This strategic plan was established based on the research objectives and experiences from published studies. Establishment standards strategy plan can ensure the researchers to implement the effective, standard, and usable national registry.

Designing the framework for a different level of data

The survey indicated that one of the most important challenges in twin registries is data gathering. Without the standard structure of data sets, gathering information will be useless. Three proposed main categories for gathering data in Iran are twin newborns, school-aged twins, and adults. This classification facilitates the gathering, storing, and analyzing twin's data. Moreover, the source of data can be defined easier.

Based on the analysis of other twin registries' achievements and FG discussions, the suggested framework was designed for the Iranian twin registry. The proposed hierarchical model of different data levels in the twin registry is shown in Fig. 3.

The data of twins can be obtained from three levels of data. In the lowest level, raw data can be collected from distributed resources which include all of the studies and researches have been conducted in twins. Aggregating these raw data is very difficult because they did not have the same structure. Additionally, the distributed nature of these kinds of data is another challenge. If such information is to be used, it must be cleaned and then aggregated according to the standard data format of the registry.

In the next level, twin's data could be obtained from different platforms. Data at this level have a more specific and defined structure based on their origin. Data can be collected directly by a voluntary collaboration of twins in different age groups through electronic or paper forms prepared by twin registry executives, or through governmental systems at the national level. These governmental and clinical resources include national infant registry, hospital systems, and ministry of education systems for gathering school-aged twins' data. The existence of a student information system in Iran can facilitate the collection of school-aged twins' data.

In the third level, after the implementation of the twin registry system, the information should be entered manually by the trained registrars in different sections based on different age groups. The data entered in this section mainly includes all of the data related to tracking twins after entering the registry. Finally, at higher levels, data is processed by aggregating data from different levels and provided to managers and researchers as information in the form of various reports for further researches.

Finally, if the system is equipped with intelligent tools such as data mining tools or artificial intelligence, knowledge can be achieved by automatic analysis of information. This knowledge is very effective in creating a hypothesis and achieving valuable results.

The proposed framework in this section is based on the study of established twin registers around the world and the analysis of the collected opinions of experts who are interested in twin studies. Recognizing the different levels of data can assist researchers to identify the appropriate data sources.

Designing a workflow model of information gathering

Reviewing the retrieved studies indicated that different methods were employed by twin registries in terms of data collection and saving data. These methods vary from country to country. It seems that the applied method depends on several factors, such as infrastructure, facilities, government policies, scope, and the main purpose of the twin registry.

The information flow for data gathering in the Iranian twin registry was designed based on our objectives, analysis of retrieved studies, and expert consensus to prevent irrelevant processes.

The optimized model was approved after an agreement between experts through several sessions of discussion and based on their feedback. The final model is represented in Fig. 4.

As it is apparent, the first step in adding twins' information to the twin registry is saving the essential and primary information. Next, after verifying the accuracy of primary information, the complementary data of twins will be added to the system database. This information can be synchronized with existing database information and added into the twin registry database. If the registry is equipped with Biobank, the biogenetic data must be linked to their original data based on the specified code of each twin.

Furthermore, we find out the mother's information could be the first step to classify twins. So, the mother's information was considered as a primary key in our database. Thus, all of the twin's data would be categorized based on mothers' data to track twins. All information defined by the three age groups can be categorized based on the mother's information of each twin. This type of categorization makes a comparison of identical and non-identical twin's data easier.

Discussion:

Due to the importance of twin studies, founded twin registries around the world were examined from the view of medical informatics. The results revealed that national twin registries are not solely restricted to a single database or data repository. Indeed, they defined as a systematic process that comprises of setting research committees, creating structured minimum datasets, designing databases, and gathering health-related information. Besides, we found that registry implementation and digitizing all twins' data at the national level could facilitate managing, analyzing, retrieving, and manipulating data for twin studies [16, 28]. Though conceptual frameworks have the potential to be used to describe the flow of information and

how to manage data, so far, no specific logical model has been proposed for the management and organizing of twin information at the national level. The authors believed that research gaps in organizing twins' health data and implementing a national twin registry have been filled by providing a logical model. Hence, the standard framework for collecting and recoding Iranian twin's health data was defined based on the combination of literature review and expert consultation.

Due to the geographical spread of the twin population and scattered data sources, data collection is one of the most challenges that all of the registries encountered. At the national level, population data sources are usually much broader. Therefore, researchers concluded that defining a suitable framework for a different level of data is essential to overcome the challenges of recording the twin's data.

To overcome the difficulty of gathering twin's data from distributed sources, conducting a parallel multidisciplinary group study, and placing twins in three different classifications is a good idea. This classification makes it possible to access information and collect data in three different groups in parallel. To our knowledge, this solution has not been used by other twin registries.

The suggested framework (Fig. 3) can be considered as a hierarchical model that was consistent with the information pyramid. As it is obvious, raw data is collected at the lowest level of the pyramid, and they transform into knowledge through the identified processes. Such a knowledge can be useful for decision-makers and health policymakers.

Furthermore, considering the mother's data as the primary key could be overcome the challenge of linking twin's data. To enhance accessibility, saving the twins' data based on the mother's information is the best strategy. This idea can be made designing the database easier too.

Regarding a platform, studies suggested that a web-based platform could be the best solution. To ensuring creating a reliable platform, registry system should be designed based on the specific requirements of each country according to its facilities, knowledge, infrastructure, and budget. Thus, it will vary the country from other countries.

In summary, we integrate all of the findings and results in overall model of Iranian twin registry [Fig. 5]. Since the main objective of a twin registry is improving twin studies, we considered the follow-up section to gather relevant health-related data continuously.

Without defining this reasonable workflow, organizing twin's data at the national level is challenging. The proposed models for the Iranian twin registry cover all twin pairs in the country with different age groups, education levels, and social status. This rational workflow was defined based on comparative studies to facilitate data gathering. One of the significant benefits of defining rational workflow is underlined in its ability to prevent entering invalid data. According to the standard workflow, data will not be recorded in the database until the demographic data are verified. Since follow-up and surveillance are so important, a special health care plan was developed to improve the health of the enrolled twins.

Through this survey, we also faced some limitations regarding access to information. Some registries have not been published their accomplishments or characteristics in the form of a scientific article or published resources. Hence, our results were completed by referring to their website, sending an email, contacting twin registries, and other unpublished resources. We should mention that the overall framework is based on the views of Iranian experts and their objectives, and it may not be generalizable to other studies.

Conclusion:

Taken together, the model proposed in this study, in addition to providing a standard method for collecting and organizing twin's data, can be considered as the first step in setting up a twin registry. Since other studies suggested that the best way to create a reliable platform is designing a system based on the requirement, it should also be a conceptual model for software development. However, the overall framework is based on the views of Iranian experts and their objectives, and it may not be generalizable to other studies.

Abbreviations

FG: Focus group

Declarations

Ethics approval and consent to participate:

The study protocol was reviewed and approved by the ethical committee at the Tehran University of medical sciences (IR.TUMS.VCR.REC.1398.128). Informed verbal consent was obtained from participants after been informed about the details of the study before focus groups and interviews.

Consent for publication

No individual details, images, or videos are included such that consent to publish is not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding:

This project was funded by Tehran University of Medical Sciences (No:39210-159-03-97). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the

Authors' contributions:

H.A and M.G contributed to the conception and design of the project. M.G extracted information of the registries and analyzed it under H.A supervision. M.G designed the models by other authors' consultation. All of the authors validated the models and frameworks. The final extracted models evaluated by H.A. H.A, L.S, M.S, and K.Z. All authors drafted the manuscript, reviewed, provided input, and approved the final manuscript.

Acknowledgments:

The authors thank all the experts of the expert panel.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

References

1. Busjahn A: **HealthTwist: the Berlin Twin Registry for health research.** *Twin Research and Human Genetics* 2013, **16**(1):163-166.
2. Botkin JR, Anderson R, Staes C, Longo N: **Developing a National Registry for conditions identifiable through newborn screening.** *Genetics In Medicine* 2009, **11**:176.
3. Cozen W, Hwang AE, Cockburn MG, Hamilton AS, Zadnick J, Mack TM: **The USC Adult Twin Cohorts: International Twin Study and California Twin Program.** *Twin research and human genetics : the official journal of the International Society for Twin Studies* 2013, **16**(1):366-370.
4. Bell JT, Spector TD: **A twin approach to unraveling epigenetics.** *Trends in Genetics* 2011, **27**(3):116-125.
5. Busjahn A: **Twin registers across the globe: what's out there in 2002?** *Twin research : the official journal of the International Society for Twin Studies* 2002, **5**(5):v-vi.
6. Wood AC, Neale MC: **Twin studies and their implications for molecular genetic studies: endophenotypes integrate quantitative and molecular genetics in ADHD research.** *Journal of the American Academy of Child & Adolescent Psychiatry* 2010, **49**(9):874-883.
7. Craig JM: **Epigenetics in Twin Studies.** *Medical Epigenetics* 2013, **1**(1):78-87.
8. Sahu M, Prasuna JG: **Twin studies: A unique epidemiological tool.** *Indian J Community Med*, 2016, **41**(3):177.

9. Hickey GL, Grant SW, Cosgriff R, Dimarakis I, Pagano D, Kappetein AP, Bridgewater B: **Clinical registries: governance, management, analysis and applications.** *European Journal of Cardio-Thoracic Surgery* 2013, **44**(4):605-614.
10. Kotoulas A, Lambrou G, Koutsouris D-D: **Design and virtual implementation of a biomedical registry framework for the enhancement of clinical trials: colorectal cancer example.** *BMJ Health & Care Informatics* 2019, **26**(1):e100008.
11. Bellgard MI, Macgregor A, Janon F, Harvey A, O'Leary P, Hunter A, Dawkins H: **A modular approach to disease registry design: Successful adoption of an internet-based rare disease registry.** *Human Mutation* 2012, **33**(10):E2356-E2366.
12. Tiro JA, Kamineni A, Levin TR, Zheng Y, Schottinger JS, Rutter CM, Corley DA, Skinner CS, Chubak J, Doubeni CA *et al*: **The colorectal cancer screening process in community settings: a conceptual model for the population-based research optimizing screening through personalized regimens consortium.** *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2014, **23**(7):1147-1158.
13. Hoque DME, Kumari V, Hoque M, Ruseckaite R, Romero L, Evans SM: **Impact of clinical registries on quality of patient care and clinical outcomes: A systematic review.** *PLOS ONE* 2017, **12**(9):e0183667.
14. Bellgard MI, Napier KR, Bittles AH, Szer J, Fletcher S, Zeps N, Hunter AA, Goldblatt J: **Design of a framework for the deployment of collaborative independent rare disease-centric registries: Gaucher disease registry model.** *Blood Cells, Molecules, and Diseases* 2018, **68**:232-238.
15. Krasnow RE, Jack LM, Lessov-Schlaggar CN, Bergen AW, Swan GE: **The twin research registry at SRI international.** *Twin Research and Human Genetics* 2013, **16**(1):463-470.
16. Hur Y-M, Craig JM: **Twin Registries Worldwide: An Important Resource for Scientific Research.** *Twin Research and Human Genetics* 2013, **16**(1):1-12.
17. Castillo-Fernandez JE, Spector TD, Bell JT: **Epigenetics of discordant monozygotic twins: implications for disease.** *Genome Medicine* 2014, **6**(7):60.
18. Malki K, Koritskaya E, Harris F, Bryson K, Herbster M, Tosto MG: **Epigenetic differences in monozygotic twins discordant for major depressive disorder.** *Translational Psychiatry* 2016, **6**(6):e839-e839.
19. Chulada P, Corey L, Vannapagari V, Whitehead N, Blackshear P: **The Feasibility of Creating a Population-based National Twin Registry in the United States.** *Twin Res Hum Genet* 2006, **9**(6): 919-926.
20. Morsy A, Lim TO, Varatharajan S, Jie Ying L: **National registries in developing countries: Understanding construction challenges and implementation steps.** In: *2010 5th Cairo International Biomedical Engineering Conference: 16-18 Dec. 2010* 2010; 2010: 75-78.
21. Braun V, Clarke V: **Using thematic analysis in psychology.** *Qualitative Research in Psychology* 2006, **3**(2):77-101.

22. Skytthe A, Kyvik KO, Holm NV, Christensen K: **The Danish Twin Registry.** *Scand J Public Health* 2011, **39**(7 Suppl):75-78.
23. Hopper JL, Foley DL, White PA, Pollaers V: **Australian Twin Registry: 30 years of progress.** *Twin research and human genetics : the official journal of the International Society for Twin Studies* 2013, **16**(1):34-42.
24. Magnusson PK, Almqvist C, Rahman I, Ganna A, Viktorin A, Walum H, Halldner L, Lundstrom S, Ullen F, Langstrom N *et al*: **The Swedish Twin Registry: establishment of a biobank and other recent developments.** *Twin research and human genetics : the official journal of the International Society for Twin Studies* 2013, **16**(1):317-329.
25. Willemsen G, Vink JM, Abdellaoui A, den Braber A, van Beek JH, Draisma HH, van Dongen J, van't Ent D, Geels LM, van Lien R: **The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection.** *Twin Research and Human Genetics* 2013, **16**(1):271-281.
26. Moayyeri A, Hammond CJ, Hart DJ, Spector TD: **The UK Adult Twin Registry (TwinsUK Resource).** *Twin Research and Human Genetics* 2012, **16**(1):144-149.
27. Kaprio J: **The Finnish Twin Cohort Study: an update.** *Twin research and human genetics : the official journal of the International Society for Twin Studies* 2013, **16**(1):157-162.
28. Buchwald D, Kaprio J, Hopper JL, Sung J, Goldberg J, Fortier I, Busjhan A, Sumathipala A, Cozen W, Mack T: **International Network of Twin Registries (INTR): Building a platform for international collaboration.** *Twin Research and Human Genetics* 2014, **17**(6):574-577.
29. Cotichini R, Fagnani C, Patriarca V, Nistico L, Brescianini S, Cirrincione R, D'Ippolito C, Pulciani S, Figa-Talamanca L, Cannoni S *et al*: **Twins in biomedical research and the creation of the "National Twin Registry".** *Epidemiol Prev* 2003, **27**(5):297-302.
30. Croft ML, Read AW, de Klerk N, Hansen J, Kurinczuk JJ: **Population based ascertainment of twins and their siblings, born in Western Australia 1980 to 1992, through the construction and validation of a maternally linked database of siblings.** *Twin research : the official journal of the International Society for Twin Studies* 2002, **5**(5):317-323.

Figures

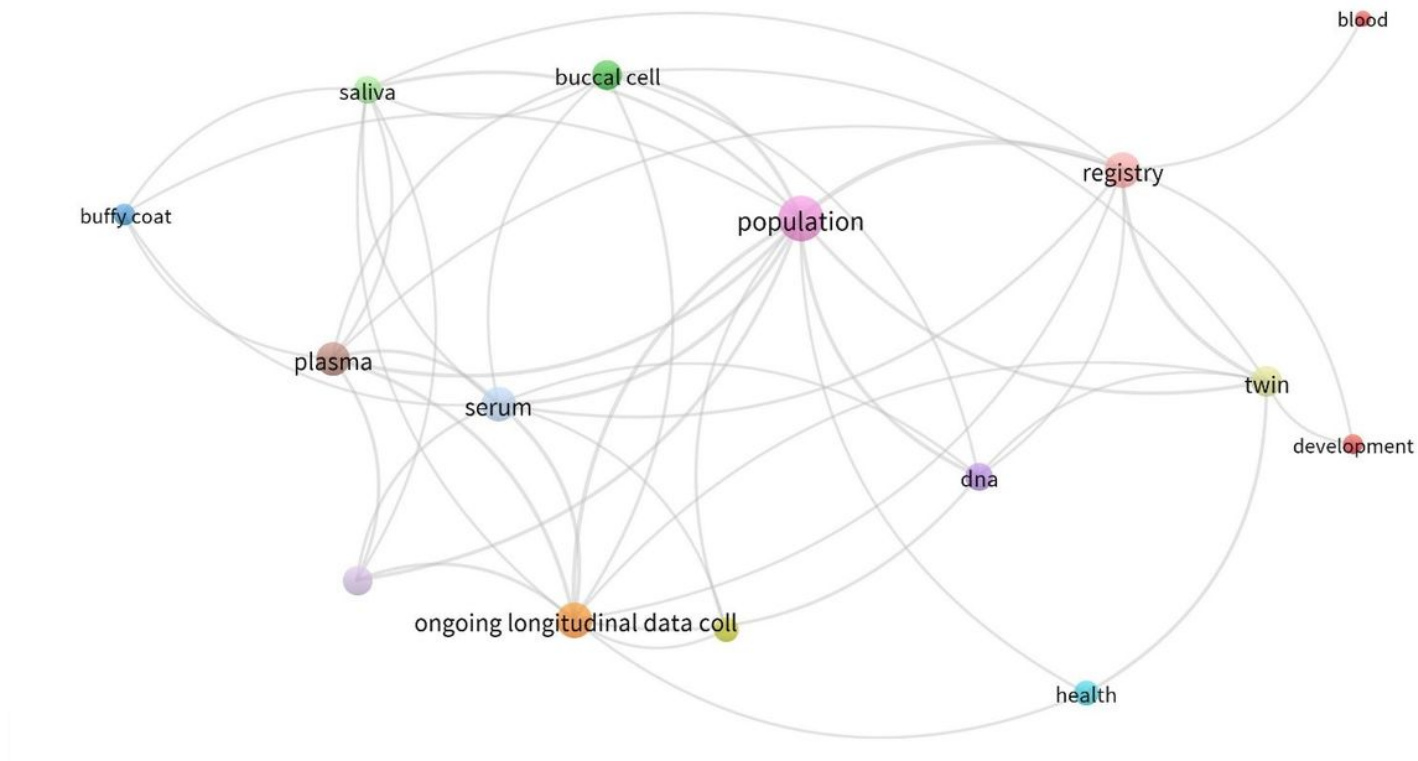


Figure 1

The network of the most frequent words from the analysis of results

Step 1:	Identifying the steering committee
Step 2:	Describing the final goals and objectives of the twin registry
Step 3:	Specifying the conceptual model of twin registry and development process
Step 4:	Defining the workflow and the data collection methods
Step 5:	Determining the data set elements and producing data dictionary based on the expert consensus
Step 6:	Developing and implementating the software
Step7:	Collecting the complementary data and developing biobank
Step 8:	Analyzing, reporting, and classifying the obtained data for further research
Step 9	Survivelance

Figure 2

The different steps of the strategy plan

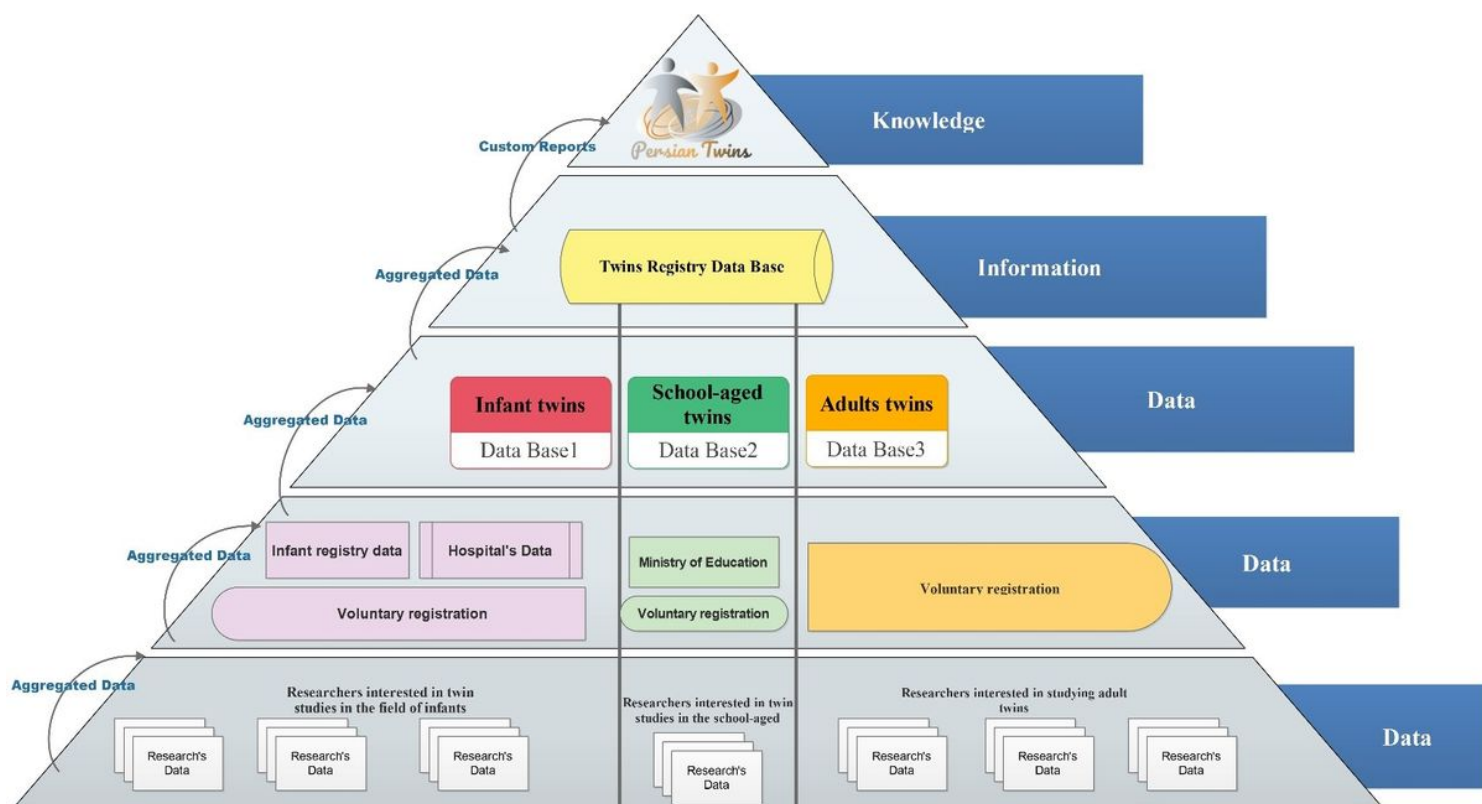


Figure 3

The suggested framework of Iranian twin registry

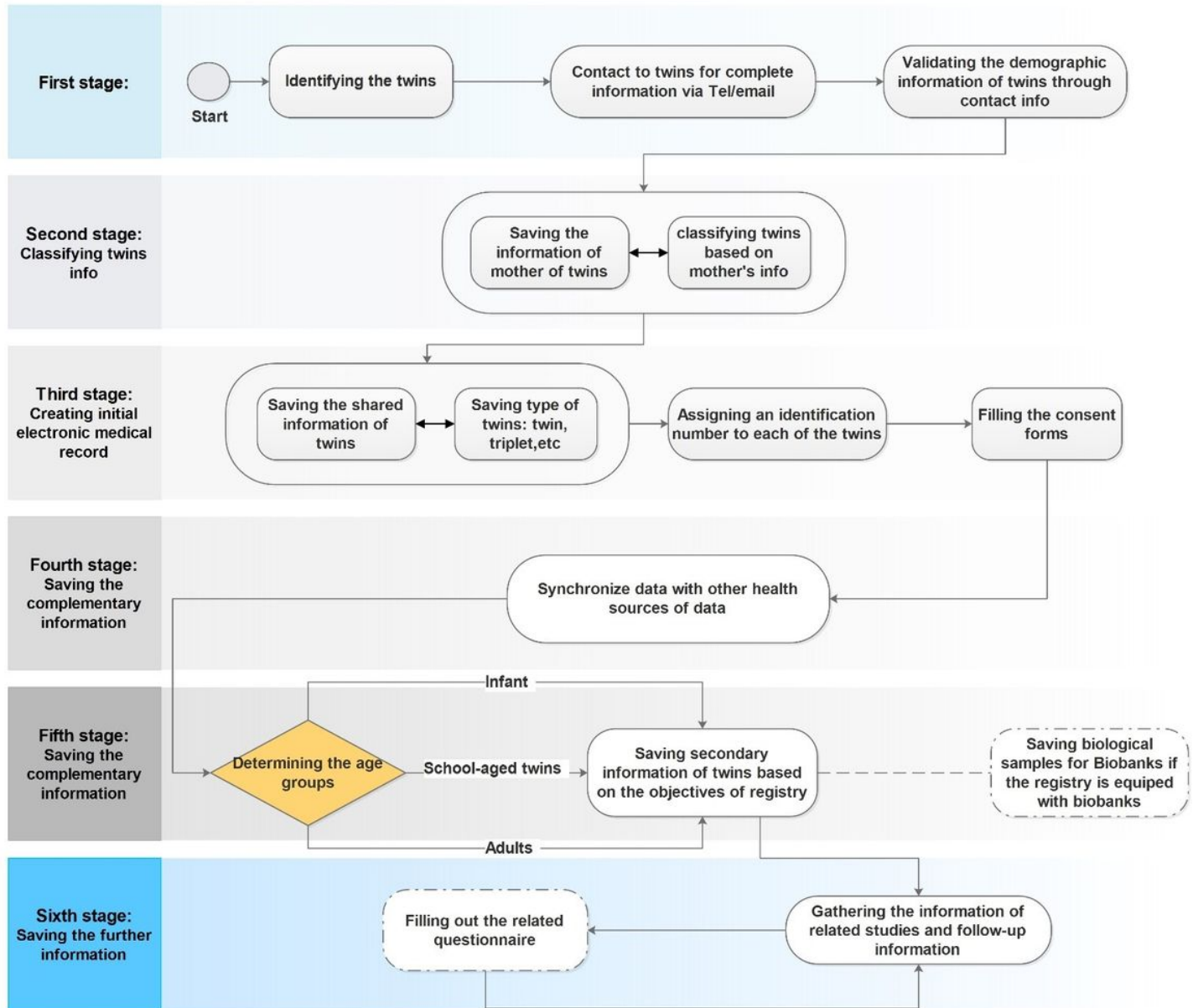


Figure 4

The suggested workflow model for gathering data of Iranian twins

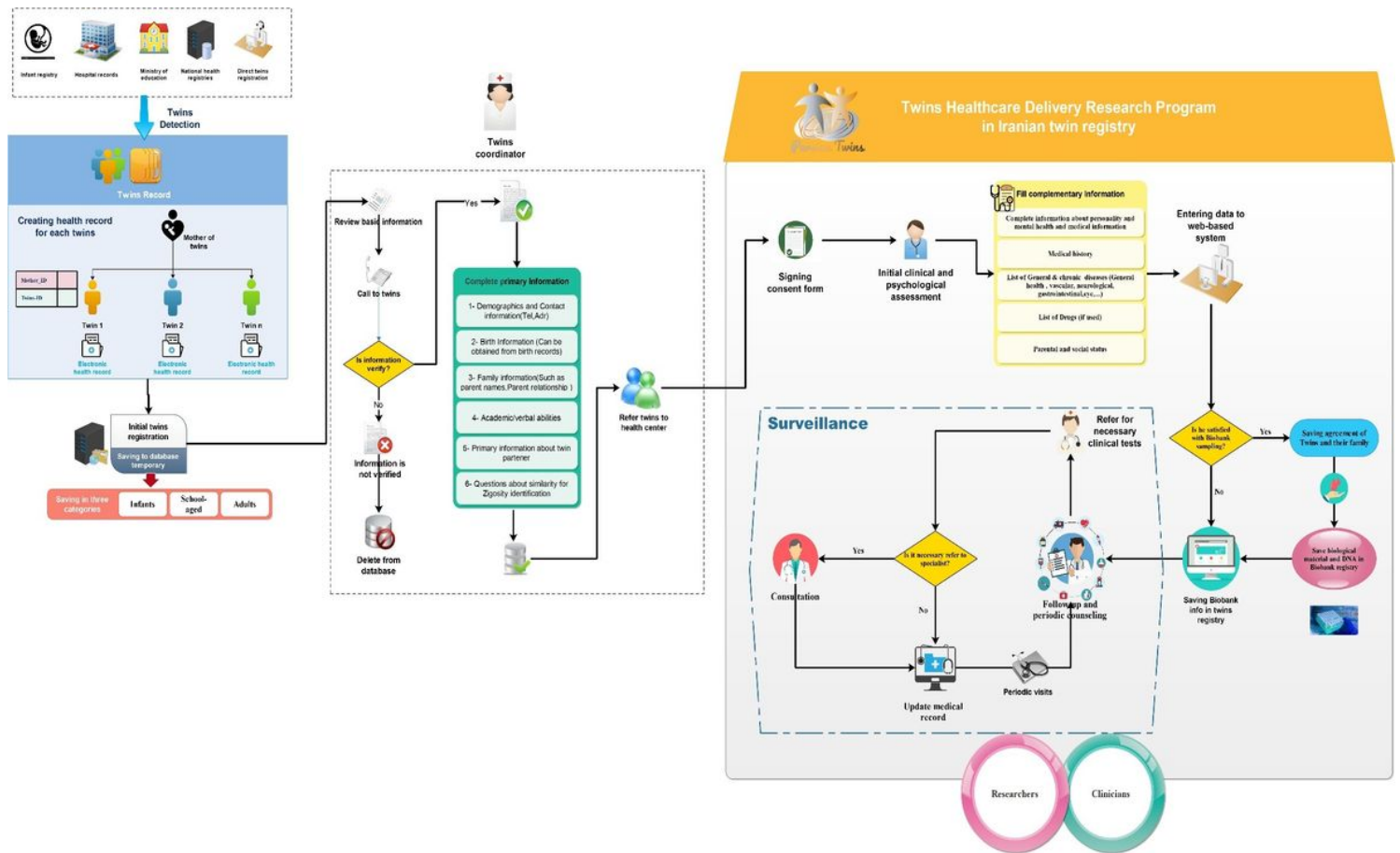


Figure 5

The schematic model of national Iranian twin registry

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SUPPLEMENTCOREQchecklisttwinregistry.docx](#)