



# The Research Attitude Required for Engineers Who are Involved in the Development of Data Warehouses for Epidemiological Analysis: Let us Provide Value, not Just a Labor Force

**Tomohide Iwao\***

*Institute for Research Initiatives, Nara Institute of Science and Technology, Japan*

**Received:**  November 24, 2018; **Published:**  November 30, 2018

**\*Corresponding author:** Tomohide Iwao, Institute for Research Initiatives, Nara Institute of Science and Technology, Japan

## Opinion

Recently, especially in Asian countries such as Korea, Taiwan, and Japan, databases that store treatment information used for insurance claims (insurance claims databases) have been set up at the national level. Some of these insurance claims databases have also been made publicly available for secondary purposes such as research and studies, and many research papers have already been published especially in the field of epidemiology. However, these databases were not constructed with the expectation of such secondary uses, so there are many cases where the data required for analysis is insufficient. The most well-known databases use the relational model that is based on predicate calculus and set theory [1]. The relational model is used for each country's insurance claims database [2] and normalization is performed as is appropriate for data storage and update [3]. Therefore, data related to a patient is distributed across multiple tables. As a result, epidemiological research studies that use these databases (database research), require special knowledge concerning medical and statistical analysis as well as techniques for reconstructing database operations on the datasets for each patient while supplementing them with the necessary data for epidemiological research. In Japan, the database called the NDB (National Database of Health Insurance Claims and Specific Health Checkups of Japan) is run by the Ministry of Health, Labor, and Welfare [4]. The authors are researchers in the engineering field who have been engaging in such research since 2015, when the trial operation of the NDB began. Since that time, there have been almost no other researchers from the engineering field who have participated in NDB, even inside academia and the government. This thesis analyzes the

background and reasons for this situation, with the purpose of describing the role that researchers in the engineering field should play in database research.

## Characteristics of Database Research

The main classifications of medical research are basic research in biological and medical fields, database research and clinical research. The comparison of database research and basic research is interesting and the two have different characteristics. Additionally, the research processes that require labor are different in both classifications. In general, for basic research, a large amount of labor is required for experimentation and data gathering, while for database research, little time is required for data gathering because data is reused. On the other hand, while database research or basic research, is often the type of research that validates hypotheses, exploratory data analysis in order to construct a hypothesis from the data is important [5]. Therefore, the data analysis (data handling) load, including the database operations mentioned above, is usually large. In other words, in basic research, labor is required mostly for the experimentation and data gathering processes, while in database research, labor is required mostly for data analysis.

However, a very important research process for database research is the construction of a hypothesis. So, if a researcher who presents a hypothesis and a researcher who performs the data analysis write a paper together, the former is often listed as the first author. The perception that only the construction of the hypothesis requires brain work has been inherited from basic research and is deeply rooted.

## Characteristics of Researchers in Epidemiology Who Perform Database Research

In Japan in particular, researchers in epidemiology who have a medical license have a very high social status. They are also often involved in a wide range of activities including basic or database research and treatment making them very busy. There are only a few researchers who are specialists in database research. Yet the support of a researcher who is familiar with database technology is often required in the engineering field. In addition, research speed is emphasized, because researchers in epidemiology want to obtain research results quickly. When researchers in the engineering field who provide such support are not actively incorporated into the epidemiological research they automatically become "helpers" and merely provide labor.

In the database research projects, I have been involved in so far, there were some researchers in epidemiology who had medical licenses and were well versed in programming techniques and database technology. However, they were only able to solve their research problems independently. With regard to the know-how that has been received from these researchers in the research process thus far, source code has been given, but it could only be used under limited conditions and was not user friendly. Based on a little imagination, the reason for this is simple. These researchers may have good skills such as programming, but such skills are not their regular job, and they were satisfied just to display their multiple talents, such as in engineering technology. They knew that the development of generic systems was not their role. In this sense, having good skills does not necessarily result in developing an excellent system.

## The Role that Researchers in the Engineering Field Should Play in Database Research

In the basic research fields of medicine and biology, the development of experimental methods and experimental equipment for the purpose of verifying hypotheses effectively is considered important. The inventors of these experimental methods and equipment are often researchers and developers that have doctorate degrees in their fields. This concept should also apply to database research. Researchers in the engineering field who are involved with database research should come up with their own research questions and gain a lot of experience in epidemiological research.

ISSN: 2574-1241

DOI: [10.26717/BJSTR.2018.11.002123](https://doi.org/10.26717/BJSTR.2018.11.002123)

Tomohide Iwao. Biomed J Sci & Tech Res



This work is licensed under Creative Commons Attribution 4.0 License

Submission Link: <https://biomedres.us/submit-manuscript.php>

This is because only through understanding epidemiological research, can one identify research problems for database research. It would also be ideal for researchers in epidemiology, including clinicians, to develop systems such as data warehouses that enable them to conduct research independently.

## Conclusion

This thesis has discussed the problem of database research and the role that researchers in the engineering field should play in it. In database research, as know-how, the creation of source code and manuals is useful to a certain degree, but not decisive in effectiveness. In order to use them, a certain amount of basic knowledge and time are required, so it is not uncommon for them to barely be used. Accordingly, it is appropriate to develop systems (software) with excellent interfaces that can be used even by beginners without requiring basic knowledge and time. It is difficult to develop a system that can cover all databases, but even just developing a system limited to the use of one database would be of immense value. In database research, from the point of view of a researcher in epidemiology, a researcher in the engineering field may often be seen as a laborer in charge of data handling. The most important thing for a researcher in the engineering field is to have a research attitude for inventing a system that will be useful someday, while providing labor.

## Acknowledgement

This article was performed with the assistance of a research grant from the Ministry of Education, Culture, Sports, Science, and Technology (Number: 17K1781600).

## References

1. EF Codd (1970) A relational model of data for large shared data banks. Communications of the ACM 13(6): 377-387.
2. Logyoung Kim, Jee-Ae Kim, Sanghyun Kim (2014) A guide for the utilization of Health Insurance Review and Assessment Service National Patient Samples. Epidemiology and Health 36: e2014008.
3. CJ Date (2005) Database in Depth: Relational Model for Practitioners. O'Reilly & Associates Inc pp. 136-138.
4. (2018) Ministry of Health, Labor and Welfare. Website on the provision of healthcare claims data and data from specific health examinations.
5. Sergeant Evan, Perkins Nigel (2015) Epidemiology for Field Veterinarians: An Introduction. Cab Intl pp. 156-177.



## Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

<https://biomedres.us/>