Current Topics

Drug Discovery: Recent Progress and the Future

Review

Drug Repositioning and Target Finding Based on Clinical Evidence

Shuji Kaneko* and Takuya Nagashima

Department of Molecular Pharmacology, Graduate School of Pharmaceutical Sciences, Kyoto University; 46–29 Yoshida-Shimoadachi-cho, Sakyo-ku, Kyoto 606–8501, Japan.

Recent pharmacological studies have been developed based on finding new disease-related genes, accompanied by the production of gene-manipulated disease model animals and high-affinity ligands for the target proteins. However, the emergence of this gene-based strategy in drug development has led to the rapid depletion of drug target molecules. To overcome this, we have attempted to utilize clinical big data to explore a novel and unexpected hypothesis of drug–drug interaction that would lead to drug repositioning. Here, we introduce our data-driven approach in which adverse event self-reports are statistically analyzed and compared in order to find and validate new drug targets. The hypotheses provided by such a data-driven approach will likely impact the style of future drug development and pharmaceutical study.

Key words   data-driven study; clinical big data; adverse event reporting system; target finding

1. INTRODUCTION

In Japan, medical care is one of the industries delayed in utilizing computerization. The term 'computerization' here means using data in various ways, such as accumulating data in a machine-readable, normalized form, and objectively optimizing practices through statistical analyses of big data. From a traditional point of view, medical information should be handled carefully to protect individual privacy; for that reason, valuable clinical data has long remained in-house. However, the sharing of medical information has already been advanced in the United States and the United Kingdom, and has been introduced very recently in Japan. This was promoted by Japan's national policy for the effective use of medical information in keeping with the development of relevant laws and guidelines. From now on, it is evident that medical care will be an advanced information industry.

It is expected that treatment level and safety monitoring will be improved by shared medical information. But is there any way to utilize shared data in basic studies related to pharmaceutical sciences? In this review, based on our experiences published in 2016,[1] we show how to process clinical evidence for the promotion of basic research, and propose new methods for developing drug discovery.

2. WHAT IS U.S. FOOD AND DRUG ADMINISTRATION (FDA) ADVERSE EVENT REPORTING SYSTEM (FAERS)?

Since unfavorable side effects (hereinafter referred to as 'adverse events') of drugs are generally low in incidence, it is extremely difficult to fully grasp them at the clinical trial stage, making post-marketing surveillance obligatory. The database of adverse events is an accumulation of spontaneous reports of adverse events collected through post-marketing surveillance, to serve as valuable clinical evidence that records 'real' adverse drug effects in humans. This information resource is fully utilized for safety monitoring in regulatory science, but has not been used during basic research or the drug discovery process.

The FAERS[2] is a voluntary reporting database accumulated by the FDA since the early 2000s and released as publicly accessible data. FAERS is a relational database consisting of seven tables (Fig. 1). By the end of 2018, more than 10 million cases had been registered, and various patient symptoms can be analyzed. When analyzing FAERS data, careful consideration of such variables as the exclusion of duplicate cases, identification of drug names, and elimination of reporting bias are necessary, as described later. Nevertheless, FAERS can be used for the analysis of drug interactions and statistical studies on confounding factors of adverse events. There are over 5000 active pharmaceutical ingredients appearing in FAERS, so there are 25 million possibilities, even if simply considering a combination of two active substances, though most of the hidden possibilities have been neglected.

3. DRUG REPOSITIONING USING FAERS

Drug adverse events are often utilized in pharmacological studies to produce pathological animal models. This is because the pathology and/or phenotypes (symptoms) of adverse events reflect at least partly those seen in human disease
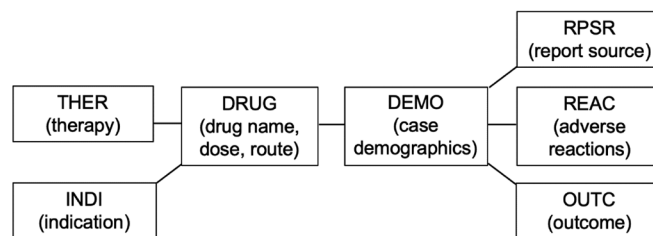

Fig. 1.   Relational Database Structure of FAERS

*To whom correspondence should be addressed.   e-mail: skaneko@pharm.kyoto-u.ac.jp
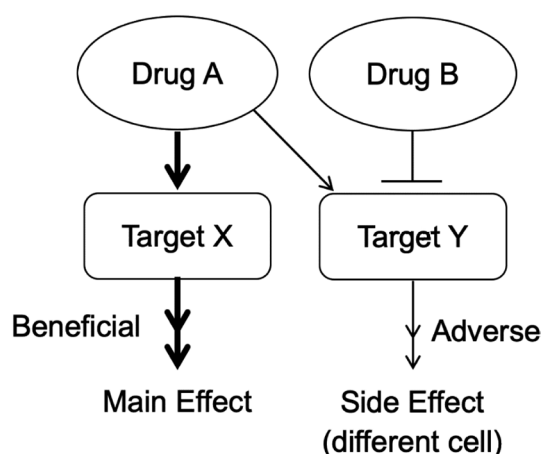
Fig. 2. A Simplified Pharmacological Concept of Drug Effects and Interactions

Table 1. The Top 30 Drugs and Reactions Reported in FAERS

Drug	Frequency	Reaction	Frequency
Adalimumab	417926	Drug ineffective	373103
Etanercept	362030	Nausea	291742
Aspirin	351357	Death	259183
Thyroxine	233750	Fatigue	238537
Atorvastatin	201407	Headache	225662
Peritoneal dialysis solution	200806	Dyspnea	205801
Acetaminophen	197452	Pain	197391
Interferon beta-1a	195996	Diarrhea	193158
Omeprazole	188852	Dizziness	192154
Furosemide	187217	Vomiting	177724
Metformin	185999	Malaise	159613
Multi-vitamin	181810	Asthenia	141247
Methotrexate	180837	Pyrexia	133226
Prednisone	179828	Rash	131796
Metoprolol	171812	Arthralgia	130929
Lenalidomide	166647	Fall	122133
Amlodipine	152457	Pain in extremity	118167
Natalizumab	151242	Myocardial infarction	112806
Infliximab	146729	Anxiety	111249
Simvastatin	146697	Injection site pain	110697
Lisinopril	144141	Insomnia	108625
Warfarin	140354	Pneumonia	107171
Esomeprazole	138412	Pruritus	107162
Quetiapine	137542	Depression	105421
Pregabalin	136143	Off label use	103488
Salbutamol	135880	Weight decreased	97936
Citalopram	132608	Feeling abnormal	95734
Gabapentin	129644	Drug dose omission	90938
Teriparatide	121470	Abdominal pain	89299
Rivaroxaban	117496	Cerebrovascular accident	87872

situations. Various pharmaceuticals have been developed by pharmacological evaluation using such disease animal models.

When drug A exerts its main action P for clinical treatment *via* biomolecule X, adverse event Q can be regarded to be mediated *via* a different biomolecule Y in other organs or cells (Fig. 2). Sometimes the main action P and the adverse event Q are caused through a common biomolecule X depending on cell-specific, different downstream signaling pathways. For example, common cyclooxygenase inhibition by aspirin causes anti-inflammatory action and gastric ulcer formation in different tissues.

Since FAERS is an accumulation of adverse events in humans, it is considered a drug-induced symptom database of humans. Thus, it is scientifically interesting to examine confounding factors that may affect the occurrence of drug-induced symptom, as long as FAERS contains many cases of polypharmacy. Therefore, if we find a case in which concomitantly-used drug B unexpectedly lowers the incidence of the drug A-induced adverse event, the drug B could be a practical remedy for the adverse event in humans, which thus enables the repositioning of an approved drug B to lower the risk of drug A. Moreover, if we identify the target Y as the site of drug interaction, the target may share a common pathway with the relevant human symptom. This may lead to the discovery of a new target for a human disease.

4. ADVANTAGES AND DISADVANTAGES OF USING FAERS

Table 1 lists the top 30 most frequently reported drugs and adverse events in FAERS through the end of 2018. There are many active ingredients classically used as pharmaceuticals; in addition, many antibody drugs and low-molecular-weight drugs that have been used more recently. On the other hand, adverse events may include such simple factors as lack of efficacy and off-label use, but there are many concepts that express symptoms and signs of humans. These are some of the varying attributes specific to spontaneous reports, and it is important to fully understand the following advantages and disadvantages of using FAERS in drug development.

4.1. Advantage (1) Anonymized clinical data with many cases and free use
Compared to traditional epidemiological studies, the great-est advantage of FAERS is that it includes real clinical data on millions of cases. It has already been anonymized, so that anyone who downloads its public data can easily start a subject of research from today. AERSMine,[3,4] an online searchable website, is also available.

4.2. Advantage (2) A wide variety of adverse events are collected from all over the world
In other countries, there are many pharmaceuticals containing active ingredients that are not approved in Japan. Because each company that sells pharmaceuticals in the United States have to report all adverse events that occur with the relevant drug, FAERS substantially covers every adverse event report from most countries, including Japan. Therefore, a wide variety of drug–drug interactions not known so far may be hidden.

4.3. Advantage (3) Easy to identify symptoms regardless of disease name
Indications and adverse events in FAERS are described by the globally unified terminology, MedDRA.[5] Unlike disease name systems such as ICD-10,[6] MedDRA widely includes subjective symptoms and signs, making it suitable for analyzing transient adverse events that do not lead to permanent lesions or disease.

4.4. Disadvantage (1) An enormous number of duplicate cases are included
Since spontaneous reports to FAERS are made not only by pharmaceutical manufacturers but also by a wide variety of individuals such as doctors, patients, and lawyers, there are

many duplicate reports in which the same case is reported. In addition, since the database is updated quarterly each year, a continuation of the same case may cause duplication. In total, it is estimated that more than 10% of the reports are duplicated, which requires that the FAERS be examined and edited manually.[7]

4.5. Disadvantage (2)   Drug names are diverse

In the 10 million cases reported in FAERS, more than 690000 drug names can be found by simply comparing the names as strings. For example, morphine is a single chemical compound without considering its salt form, but as many as 1700 variations of morphine can be found in FAERS under various trade names or with extra descriptions. The "name collation" of drug names is partly done by the FDA, but it is incomplete. Using machine learning, we have created an original drug name dictionary with 99% accuracy and coverage.[8]

4.6. Disadvantage (3)   Various biases need to be considered

Immediately after marketing of a drug, the number of adverse event reports increases because the new release of a drug initially attracts wide attention and use; the number of adverse reports tends to decrease with time. In addition, when safety information is published by regulatory authorities, the number of adverse reports on the drug or similar drugs dramatically increases. Because of this human bias,[9] the correct incidence of adverse events cannot be accurately assessed by FAERS.

Also, if there are many specific adverse event reports for a drug, other adverse event signals for the drug are weakened. Conversely, if there are many reports of certain adverse events with a specific drug, the event signal of other similar drugs is also weakened. This is called competition bias.[10]

5. COOPERATION OF DRY ANALYSIS AND WET EXPERIMENTS

A general procedure is shown in Fig. 3 according to the past report[1] as an example.

5.1. Step 1 (Dry)   Derivation of interaction hypotheses by data analysis

We conducted FAERS analysis focusing on hyperglycemia and diabetes mellitus caused by atypical schizophrenia drugs. As a result, we found that quetiapine and olanzapine increased the reporting odds ratio (ROR) for hyperglycemia/diabetes with tens of thousands of cases. The increase in ROR was robust regardless of primary disease, gender, age, and so on in stratified analysis. When we searched for a drug that reduces the ROR by calculating every case of concomitant drug used with quetiapine, we found that the ROR for quetiapine-induced hyperglycemia was reduced by about one-third with the co-treatment of vitamin D with quetiapine. We recently confirmed this hypothesis by chronologically analyzing the JMDC (formerly known as Japan Medical Data Center) Claims Database,[11] in which 40% of patients underwent health checkups yearly. In these cases, fasting blood glucose levels were elevated by quetiapine, but the increase was cancelled in patients who regularly took vitamin D.
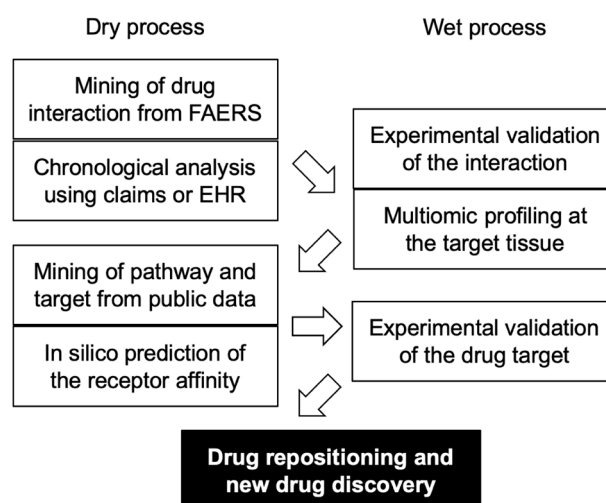

Fig. 3. Hypothesis Derivation and Demonstration Process Based on Clinical Evidence
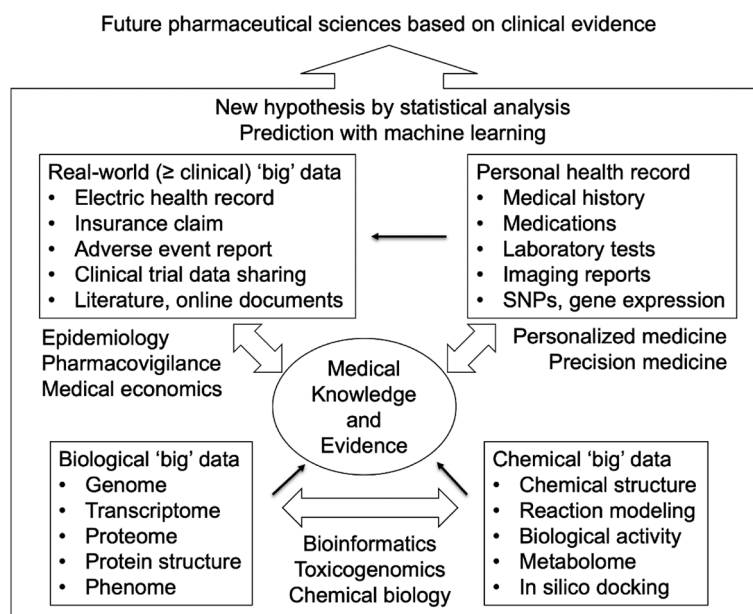

Fig. 4. Big Data and Future Pharmaceutical Sciences

5.2. Step 2 (Wet) Demonstration experiment using animals and cells

If a promising drug–drug interaction is detected by FAERS analysis, a proof experiment should be planned using experimental animals and/or primary cultured cells to confirm the hypothesis. In our example, we found that vitamin D preadministration for a week suppressed the hyperglycemia and hyperinsulinemia induced by quetiapine in mice, indicating that quetiapine caused an acute insulin resistance *in vivo*.

5.3. Step 3 (Dry) Creation of hypothesis of action mechanism using public big data

For many chemical compounds, their effects on gene expression in certain organs or cells have been comprehensively investigated and can be searched in public databases, such as the Gene Expression Omnibus. Since the involvement of insulin resistance was suggested in our study of quetiapine, we searched the toxicogenomic data for changes in the expression of genes downstream of insulin receptors, and found that the phosphatidylinositol 3-kinase (PI3K) pathway might be involved in this mechanism.

5.4. Step 4 (Wet) Demonstration of molecular mechanism

According to the hypothesis above, the expression of an adaptor protein PI3KR1 was downregulated in the skeletal muscle of quetiapine-treated mice, which was again reversed by the pretreatment with vitamin D. In addition, insulin-evoked glucose uptake into myotubes was acutely inhibited by quetiapine, which was not seen in those myotubes pretreated with vitamin D. These data indicate that vitamin D does indeed counteract quetiapine-induced insulin resistance by the upregulation of FAERS R1.

6. CONCLUSION: DATA SCIENCE FOR THE FUTURE OF PHARMACEUTICAL SCIENCES

As shown in Fig. 4, recent progress in multi-omic studies in biology, and the expansion of chemical substances with certain biological meaning, produce big data in basic sciences. The link between biological and chemical data is the molecular basis of pharmacology, and this has been enhanced by new scientific disciplines. The accumulation of big data has contributed to the enrichment of medical knowledge and evidence. By applying medical knowledge to clinical practice with genome and SNPs data, personalized or precision medicine is now being established in contemporary medicine. We are increasingly able to apply many kinds of real-world big data, such as digital health records, insurance claims and adverse event reports, in applying pharmaceuticals to human disease. A read-world database includes much more than clinical findings, as personal information and family history will ideally be included upon de-identification. These data are utilized for epidemiology, pharmacovigilance, or medical economics, but also work as a generator of hypothesis in scientific research. Although pharmaceuticals have been historically developed with translation from the animal, cell, or molecular level to clinical practice, a new "data-driven" approach based on clinical evidence will pave the way to the new era of pharmaceutical sciences. This review was written based on the Japanese article by Kaneko S. in *Farumashia*, 55, 208–212, 2019 (Seminar).

Conflict of Interest The authors declare no conflict of interest.

REFERENCES

1) Nagashima T, Shirakawa H, Nakagawa T, Kaneko S. Prevention of antipsychotic-induced hyperglycaemia by vitamin D: a data mining prediction followed by experimental exploration of the molecular mechanism. *Sci. Rep.*, **6**, 26375 (2016).
2) "FDA Adverse Event Reporting System (FAERS).": ‹https://www.fda.gov/drugs/informationondrugs/ucm135151.htm›
3) Sarangdhar M, Tabar S, Schmidt C, Kushwaha A, Shah K, Dahlquist JE, Jegga AG, Aronow BJ. Data mining differential clinical outcomes associated with drug regimens using adverse event reporting data. *Nat. Biotechnol.*, **34**, 697–700 (2016).
4) "AERS*Mine*.": ‹https://research.cchmc.org/aers/›
5) "MedDRA.": ‹https://www.meddra.org/how-to-use/support-documentation/japanese›
6) "ICD-10.": ‹http://www.dis.h.u-tokyo.ac.jp/byomei/icd10/index.html›
7) Banda JM, Evans L, Vanguri RS, Tatonetti NP, Ryan PB, Shah NH. A curated and standardized adverse drug event resource to accelerate drug safety research. *Sci. Data*, **3**, 160026 (2016).
8) Life Science Dictionary Project. ‹https://lsd-project.jp/›
9) Weiss-Smith S, Deshpande G, Chung S, Gogolak V. The FDA drug safety surveillance program: adverse event reporting trends. *Arch. Intern. Med.*, **171**, 591–593 (2011).
10) Raschi E, Poluzzi E, Koci A, Antonazzo IC, Marchesini G, DePonti F. Dipeptidyl peptidase-4 inhibitors and heart failure: analysis of spontaneous reports submitted to the FDA adverse event reporting system. *Nutr. Metab. Cardiovasc. Dis.*, **26**, 380–386 (2016).
11) Claims Database JMDC. ‹https://www.jmdc.co.jp/pharma/database.html›