Association between KEGG Biological Pathways and Adverse Drug Reactions of HIV, TB and Other Drugs Frequently Implicated in ADRs

Wilbert Sibanda
School of Nursing and Public Health, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa
Email: Sibandaw@ukzn.ac.za

***Abstract*—In South Africa, the proportion of patients admitted to hospitals with Adverse Drug Reactions (ADRs) range between 2-21.4% and between 1.7 to 25.1% of hospital in-patients are reported to have developed ADRs in hospitals. Drugs are therefore responsible for significant mobility and mortality amongst people in South Africa. The use and uptake of medicines results in ADRs due to medicine's toxicity and interactions with other medicines. A large number of ADRs are preventable. This research aims to investigate the association between Kyoto Encyclopedia of Genes and Genomes (KEGG) biological pathways involved in cellular response to drugs and occurrence of Adverse Drug Reactions (ADRs). A Pearson's Product-Moment Correlation (PPMC) was run to assess the relationship between the frequency of ADRs and biological pathways for 33 drugs that included 18 HIV drugs commonly called Anti-Retrovirals (ARVs), 5 anti-tuberculosis (TB) drugs and 10 drugs frequently implicated in ADRs (DFAs) comprising of 3 opioids, 2 loop diuretics, 2 beta agonists, 2 low molecular heparins and 1 systemic corticoid. The ratios of biological pathways/ADRs for ARVs, TB drugs and DFAs were found to be 0.02, 0.33 and 0.06 respectively. ARVs had on average more ADRs (165.22±3.94) compared to TB drugs (67.60±29.17) and DFAs (160.9±49.99). However, TB drugs were linked to a comparatively larger number of KEGG biological pathways (22±9.95) compared to ARVs (3.94± 0.74) and DFAs (9.50±1.79). Further research is required to understand the importance of these research findings towards the development of more effective drugs characterized by reduced prevalence of ADRs.**

***Index Terms*—KEGG, biological pathways, ADRs, HIV, TB**

I. Introduction

Recently, significant research has been dedicated towards understanding the involvement of biological pathways in how cells respond to drugs. Research has demonstrated the importance of investigating biological processes responsible for ADRs in the development of more effective drugs [1]. Silberberg *et al.*, emphasized the importance of understanding drug-induced signaling pathways in order to fully appreciate the modes of action and the resultant ADRs of medicinal products [2].

Main factors play an important role in the development ADRs, some of these are patient related, drug related or socially related factors. Understanding the different effects of these factors on ADRs enables healthcare professionals to select the most appropriate medication for an individual [3].

Adverse Drug Reactions (ADRs) are defined as a response to a drug which is noxious and unintended, and which occurs at doses normally used in man for the prophylaxis, diagnosis or therapy of disease, or the modification of physiological diagnosis [4]. ADRs are undesirable effects, reasonably associated with the use of the drug that may occur as part of the pharmacological action of a drug or may be unpredictable in their occurrence [5].

KEGG (Kyoto Encyclopedia of Genes and Genomes) is a collection of databases dealing with genomes, biological pathways, diseases, drugs, and chemical substances. KEGG is utilized for bioinformatics research and education, including data analysis in genomics, metagenomics, metabolomics and other omics studies, modeling and simulation in systems biology, and translational research in drug development. The KEGG database started in 1995 by Professor Minoru Kanehisa at the Institute for Chemical Research, Kyoto University [6].

KEGG Pathway mapping is the process of mapping molecular datasets, especially large-scale datasets in genomics, transcriptomics, proteomics, and metabolomics, to the KEGG pathway maps for biological interpretation of higher-level systemic functions.

Biological pathways represent a series of actions among molecules in a cell that leads to a certain product or a change in a cell. Therefore pathways have the potential to stimulate the assembly of new molecules, such as a fat or protein as well as turning on and off of genes. These biological pathways control how the body processes drugs and the most common types of biological pathways are metabolic, genetic and signal transduction pathways.

In this paper, the aim is to describe the association between ADRs and biological pathways of ARVs, TB drugs and other drugs most frequently implicated in causing ADRs (DFAs).

Manuscript received February 24, 2016; revised July 22, 2016.

doi: 10.18178/ijpmbs.5.4.227-231

II. Research Methodology

A. Data Sources

To conduct the study, data from SIDER [7] was used. SIDER presents an aggregate of dispersed public information on drug side effects and indications.

Using SIDER™, 33 drugs and corresponding 1 448 side-effects were collected. The dataset was constructed based on the approach in Liu *et al.* [8]. The 33 drugs had the following profiles, a total of 1 448 ADRs extracted from SIDER and a total of 121 biological pathways extracted from DrugBank™ [9].

Despite that each drug studied had a relatively large number of ADRs on average (43.88), the observed biological pathways linked to each drug were comparatively small (3.67). TB drugs, ethionamide and pyrazinamide were associated with the largest number of biological pathways of 49 and 42 respectively.

B. Data Analysis

A Pearson's Product-Moment Correlation (PPMC) was run to assess the relationship between the numbers of ADRs and biological pathways for all the 33 drugs studied that included 18 ARVs, 5 anti-TB drugs and 10 drugs frequently implicated in ADRs comprising 3 opioids, 2 loop diuretics, 2 beta agonists, 2 low molecular heparins and 1 systemic corticoid.

PPMC is a measure of the correlation between sets of data. It shows the linear relationship between two sets of data. The formula for PPMC r is;

$$r = \frac{\Sigma(xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y^2 - (\Sigma y)^2]}} \quad (1)$$

where:

$\{x_1 \ldots x_n\}$ in one dataset containing n values,

$\{y_1 \ldots y_n\}$ is another dataset containing n values

Table I provides a guideline for the application of Pearson's correlation coefficients based on Cohen's (1988) conventions for the interpretation of effect size [10].

TABLE I. Guidelines for Interpretation of Correlation Coefficient

Strength of Association		Coefficient, r	
		Positive	Negative
Small		0.1 to 0.3	-0.1 to -0.3
Medium		0.3 to 0.5	-0.3 to -0.5
Large		0.5 to 1.0	-0.5 to -1.0

PPMC is one of the most widely used methods for studying relationship between inter-related variables. However, it has the following limitations;

(i) Assumes a linear relationship between the variables even though it may not be there.

(ii) A high correlation, does not necessarily mean very close relationship between the variables.

(iii) PPMC is exceedingly affected by extreme values.

III. Results

Fig. 1 and Fig. 2, show the distribution of ADRs and biological pathways respectively, associated with all the 33 drugs studied in this paper.

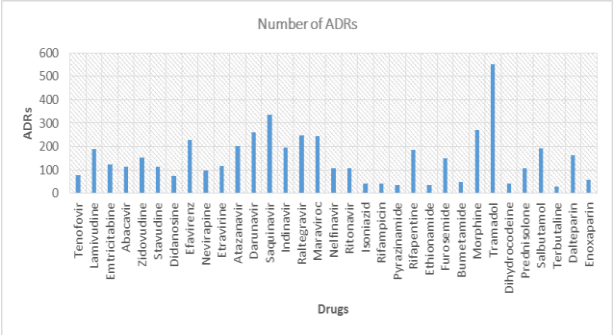

Figure 1. The ADR distribution of drugs

A close visual examination of Fig. 1 and Fig. 2, show that Tramadol has the largest number of ADRs (553), though it was linked to only 17 biological pathways. The top 5 drugs linked to the largest number of ADRs, had biological pathways ranging from 1 to 17.

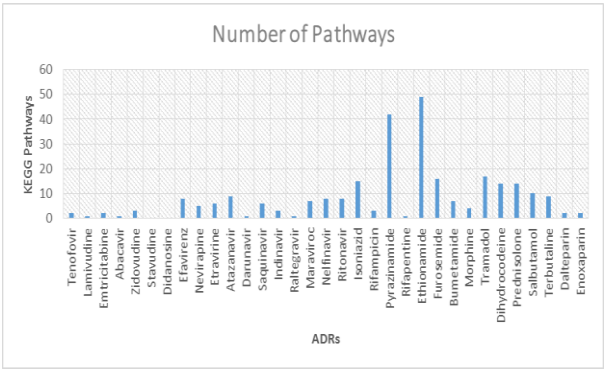

Figure 2. The pathway distribution of drugs

A. Descriptive Statistics of ADRs and Pathways

Table II shows that there is a wide difference in the numbers of ADRs between the drug that had the least number of ADRs and the one with the highest number of ADRs. The least number of ADRs for a drug was 27, corresponding to Terbutaline compared to Tramadol that had the highest number of ADRs at 553. The average number of ADRs for the 33 drugs was 149.12. The average number of biological pathways for a drug was found to be 8.36.

TABLE II. Descriptive Statistics of ADRs and Biological Pathways for All the 33 Drugs Studied

	Range	Min.	Max.	Mean	Std.Dvn	Variance
No. of ADRs	526	27	553	149.12±18.86	108.36	11740.99
No. Pathways	49	0	49	8.36±1.88	10.80	116.61

B. Pearson's Correlation (r): ADRs vs Pathways for All 33 Drugs Studied

The purpose of correlation analysis is to measure and interpret the strength of a linear relationship between two continuous variables, namely ADRs and biological pathways. Correlation coefficients take on values between -1 and +1, ranging from negatively correlated (-1) to uncorrelated (0) to positively correlated (+1).

However, high correlation between the two variables, does not necessarily imply causality.

The null hypothesis for the Pearson's correlation is that r = 0, meaning that there is no relationship between the biological pathways and ADRs.

As shown in Table III, the Pearson correlation coefficient (r) is -0.18 indicating a negative, albeit small association between the variables.

TABLE III. PEARSON'S CORRELATION BETWEEN ADRS AND BIOLOGICAL PATHWAYS FOR ALL THE 33 DRUGS STUDIED

		No. of ADRs	No. of Pathways
No. of ADRs	Pearson correlation (r)	1	-0.180
	N	33	33
No. of KBPs	Pearson correlation (r)	-0.180	1
	N	33	

TABLE IV. T-TEST: STATISTICAL SIGNIFICANCE OF THE DIFFERENCE BETWEEN BIOLOGICAL PATHWAYS AND ADRS OF ALL 33 DRUGS STUDIED

	Paired Differences				t	df	Sig. (2 tailed
	Mean	Std. Error of Mean	95% Confidence Interval				
			Lower	Upper			
No. of ADRs-Pathways	140.8	19.29	101	180	7.3	32	0.00

A t-test (Table IV), was used to test whether or not the difference between ADRs and biological pathways was significant. For the t-test result to be significant, the difference between biological pathways and ADRs should not have occurred as result of an atypical sampling.

The null hypothesis for the t-test is that there is no association between the numbers of biological pathways and ADRs across all the 33 drugs studied in this paper. The P-value was found to be 0.00, which is less than $\alpha = 0.05$. The null hypothesis was rejected, in favor of the alternative hypothesis that there was a significant difference between the two variables.

C. *Pearson's Correlation (r): ADRs vs Biological Pathways for ARVs*

Table V, indicates that the average number of ADRs for ARVs was 165.22, which was significantly higher than the average number of ADRs for all the 33 drugs studied in this research paper.

TABLE V. DESCRIPTIVE STATISTICS OF ARV DRUGS

	Drug	**ADRs**	**Pathways**
N	18		
Mean		165.22	3.94
Std. Error of Mean		17.57	0.74
Median		138	3.00
Mode		74	1
Std.Dvn.		74.54	3.15
Range		262	9
Min.		74	0
Max		336	9
Sum		2974	71

However, a smaller number of biological pathways was observed for ARVs (3.94) compared to all the 33 drugs (8.36). This means that the metabolism of ARVs is linked to a comparatively smaller number of biological pathways.

The Pearson correlation coefficient was calculated to be 0.187 indicating a small, positive association between the numbers of biological pathways and ADRs for ARVs as shown in Table VI. This means that as one variable increases, so does the other variable, albeit at a relatively smaller rate.

TABLE VI. PEARSON CORRELATION BETWEEN BIOLOGICAL PATHWAYS AND ADRS OF ARVS

		No. of ADRs	No. Biological Pathways
No. of ADRs	Pearson correlation	1	0.187
	N	18	18
No. of Biological Pathways	Pearson correlation	0.187	1
	N	18	18

On the basis of the t-test (Table VII), the P-value was found to be 0.00, which is less than $\alpha = 0.05$, resulting in the rejection of the null hypothesis. The difference in the number of biological pathways and ADRs is significant at α level of 0.05.

TABLE VII. T-TEST: STATISTICAL SIGNIFICANCE OF THE DIFFERENCE BETWEEN THE NUMBER OF BIOLOGICAL PATHWAYS AND ADRS FOR ARVS

	Paired Differences				t	df	Sig. (2 tailed
	Mean	**Std. Error of Mean**	**95% Confidence Interval**				
			Lower	**Upper**			
No. of ADRs-Pathways	161.28	17.44	124.47	198.08	9.3	17	0.00

D. *Pearson's Correlation (r): ADRs vs Biological Pathways for TB Drugs*

The average number of ADRs for the five TB drugs studied was found to be 67.60, compared to 165 ADRs for ARVs. ARVs had on average more ADRs compared to TB drugs as shown in Table VIII. However, TB drugs were found to have a significantly larger number of biological pathways (22) compared to ARVs (3.94).

TABLE VIII. DESCRIPTIVE STATISTICS OF TB DRUGS

	Drugs	**ADRs**	Biological Pathways
N	5		
Mean		67.60	22.00
Std. Error of Mean		29.17	9.95
Median		43.00	15.00
Mode		34	1
Std.Dvn.		65.23	22.25
Range		150	48
Min.		34	1
Max		184	49
Sum		338	110

Table IX, shows that the Pearson correlation coefficient was -0.583, indicating a large negative association between the numbers of biological pathways and ADRs. This means that as one variable increases the other variable decreases.

TABLE IX. PEARSON CORRELATION BETWEEN BIOLOGICAL PATHWAYS AND ADRS OF TB DRUGS

		No. of ADRs	No. of Biological Pathways
No. of ADRs	Pearson correlation	1	-0.583
	N	5	5
No. of Biological Pathways	Pearson correlation	-0.583	1
	N	5	5

TABLE X. T-TEST: STATISTICAL SIGNIFICANCE OF THE DIFFERENCE BETWEEN BIOLOGICAL PATHWAYS AND ADRS OF TB DRUGS

	Paired Differences				t	df	Sig. (2 tailed
	Mean	Std. Error of Mean	95% Confidence Interval				
			Lower	Upper			
ADRs-Pathways	45.6	35.89	-54.05	145.25	1.8	4	0.27

On the basis of the P-value of 0.273 (Table X), which was more than $\alpha = 0.05$, the test failed to reject the null hypothesis. Therefore, there was no evidence to suggest the means of biological pathways and ADRs were different at α level of 0.05. However, this does not mean the null hypothesis is true. There are numerous reasons for the failure to reject the null hypothesis, such as:

i. The null hypothesis is actually true, but may also be that,
ii. The null hypothesis is false, however there is insufficient data to provide evidence against it.

E. Pearson's Correlation (r): ADRs vs Biological Pathways for Drugs Frequently Implicated in ADRs (DFAs)

A list of drugs that are frequently implicated in ADRs was reported during a study conducted on 12 hospital wards (9 medical and 3 surgical) at the Royal Liverpool University Hospital (RLUH) over a six months period between June and December 2005 [11]. ADRs were identified in the above study based on their inclusion in either the Summary of Product Characteristics [12] and/or the British National Formulary [13].

For this research, 10 drugs were selected from the DFAs list namely 3 opioids (dihydrocodeine, morphine and tramadol), 2 loop diuretics (furosemide and bumetanide), 2 beta agonists (salbutamol and terbutaline), 2 low molecular agonists (dalteparin and enoxaparin) and 1 systemic corticoid (prednisolone).

The average number of ADRs for the 10 DFAs was 160.90, compared to 165 ADRs for ARVs and 67.60 for TB drugs (Table XI). However, DFAs, were linked on average to 9.5 biological pathways compared to 3.94 and 5 biological pathways for ARVs and TB drugs respectively.

TABLE XI. DESCRIPTIVE STATISTICS OF DFAS

	Drugs	ADRs	Biological Pathways
N	10		
Mean		160.90	9.50
Std. Error of Mean		49.99	1.79
Median		128.50	9.50
Mode		27	2
Std.Dvn.		158.10	5.66
Range		526	15
Min.		27	2
Max		553	17
Sum		1609	95

The Pearson correlation coefficient (Table XII) between number of biological pathways and ADRs was found to be 0.324, demonstrating a significant positive association between the two variables.

TABLE XII. PEARSON CORRELATION BETWEEN BIOLOGICAL PATHWAYS AND ADRS OF DFAS

		No. of ADRs	No. of Pathways
No. of ADRs	Pearson correlation	1	0.324
	N	10	10
No. of Pathways	Pearson correlation	0.324	1
	N	10	10

IV. DISCUSSION

This research clearly demonstrates that amongst the 33 drugs studied, TB drugs had the lowest average number of ADRs (67.60±29.17) compared to ARVs (165.22±17.57) and DFAs (160.90±49.99) as shown in Fig. 3 and Fig. 4.

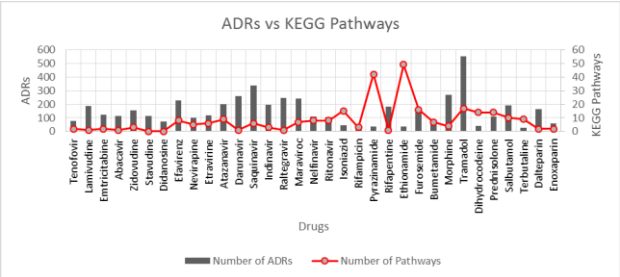

Figure 3. Biological pathways and ADRs of all the drugs in the study

The mean biological pathways/ADRs ratios for ARVs, TB drugs and DFAs were 0.02, 0.33 and 0.06 respectively as shown in Fig. 4.

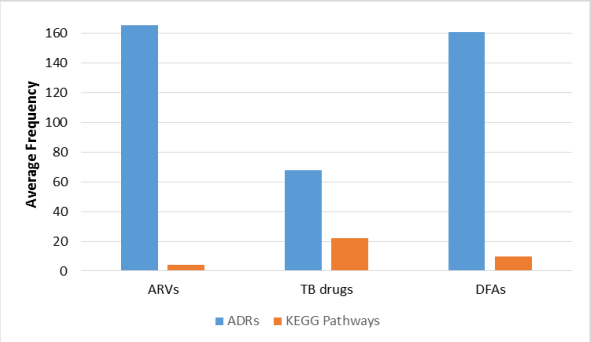

Figure 4. Average numbers of biological pathways and ADRs of ARVs, TB drugs and DFAs

The Pearson's product-moment correlation coefficients of ADRs and KEGG biological pathways for ARVs, TB drugs and DFAs were +0.187, -0.583 and +0.324 respectively. ARVs had on average more ADRs (165.22±3.94) compared to TB drugs (67.60±29.17). However, TB drugs were linked to a comparatively larger number of biological pathways (22±9.95) compared to ARVs (3.94±0.74).

Further research is required to understand the importance of these research findings towards the development of more effective drugs characterized by reduced prevalence of ADRs.

REFERENCES

[1] I. Wallach, N. Jaitly, and R. Lillian, "A structure-based approach for mapping adverse drug reactions to the perturbations of underlying biological pathways," *PLOS One*, vol. 5, no. 8, 2010.

[2] Y. Silberberg, A. Gottlieb, M. Kupiec, E. Ruppin, and R. Sharan, "Large-Scale elucidation of drug response pathways in humans," *Journal of Computational Biology*, vol. 19, no. 2, pp. 163-174, 2012.

[3] M. J. Alomar, "Factors affecting the development of adverse drug reactions," *Saudi Pharmaceutical Journal*, vol. 22, no. 2, pp. 83-94, 2014.

[4] World Health Organization (WHO), *Handbook of Resolutions and Decisions of the World Health Assembly and Executive Board WHA. Clinical Evaluation of Drugs*, 1973, vol. 11948.

[5] I. Edwards and J. Aronson, "Adverse drug reactions: Definitions, diagnosis and management," *J. Lancet*, vol. 356, pp. 1255-1259, 2000.

[6] M. Kanehisa and S. Goto, "KEGG: Kyoto encyclopedia of genes and genomes," *Nucleic Acids Res.*, vol. 28, no. 1, pp. 27-30, 2000.

[7] M. Kuhn, *et al.*, "A side-effect resource to capture phenotypic effects of drugs," *Molecular Systems Biology*, vol. 6, no. 343, 2010.

[8] M. Liu, *et al.*, "Large-Scale prediction of adverse drug reactions using chemical, biological and phenotypic properties of drugs," *J. Am. Med. Inform. Assoc*., vol. 19, pp. 28-35, 2011.

[9] C. Knox, V. Law, T. Jewson, P. Liu, and S. Ly, "DrugBank 3.0: A comprehensive resource for 'omics' research on drugs," *Nucleic Acids Res.*, vol. 39, pp. 1035-1041, 2011.

[10] J. Cohen, *Statistical Power Analysis for Behavioral Sciences*, 2nd ed., Hillsdale, 1988.

[11] E. C. Davies, *et al.*, "Adverse drug reactions in hospital in-patients: a prospective analysis of 3695 patient-episodes," *PLOS One*, vol. 4, no. 2, 2009.

[12] Association of the British Pharmaceutical Industry, *ABPI Data Sheet Compendium*, London: Datapharm Communications Ltd., 2005.

[13] British Medical Association, Royal Pharmaceutical Society of Great Britain, *British National Formulary*, London: BMA, RPS, 2005, no. 49 and 50.

Wilbert Sibanda is an associate member of the Pharmaceutical Society of South Africa. He has held the following qualifications, BSc (human physiology) (University of the Witwatersrand, Johannesburg), BSc (Med) Hons (pharmacology) (University of Capetown), MSc (Med) (pharmacy) (University of the Witwatersrand, Johannesburg), PhD (statistical modeling of antenatal HIV data) (North-West University, South Africa). He is involved in statistical modeling of HIV antenatal data in South Africa.

He is a subject specialist (bioinformatics) at the School of Information Technology, North-West University, Vaal Triangle campus, Hendrik van Eck Boulevard, Vanderbijlpark, South Africa. He has published extensively in the field of statistical modeling of HIV antenatal data.

He has chaired a special session at the International Computers and Industrial Engineering (CIE 42) Conference in Capetown, South Africa in 2012. He also chaired the Machine Learning and Databases Session at the 2013 IEEE International Conference on Bioinformatics and Biomedicine, at Tongji University, Shanghai, China. He is a reviewer of the *International Journal of Computational Biology, Informatics and Control* (IJCBIC). He is also a technical committee member of the 4th International Conference on Bioinformatics and Biomedical Science (ICBBS) to be held on the 25th to the 25th of June 2015, in Bangkok, Thailand.