2011 International Conference on Life Science and Technology
IPCBEE vol.3 (2011) © (2011) IACSIT Press, Singapore

Quantitative Validation of Clinical Studies Using Statistical Path Analysis

Application to Dopamine-Producing Paragangliomas

Smruthy K.S. , Pooja Bhat
Department of Biotechnology
PES Institute of Technology
Bangalore, India
smruthy90@gmail.com & poozabhat@yahoo.co.in

Uma S. Ranjan and Manjiri Bakre
Philips Electronics India Ltd.
Bangalore, India
uma.ranjan, manjiri.bakre@philips.com

***Abstract*—Biochemical testing is important for screening and diagnosis of several diseases. The success of biochemical tests depends on the sensitivity and specificity of metabolites for the disease. Currently, decisions regarding the abnormality of tissues or organs are made on the basis of heuristics. In case of conflicting evidence, clinicians rely on their expert judgment to arrive at a decision. Hence, it is necessary to have a system that can help clinicians make decisions regarding causes of observed conditions, quantitatively validate their heuristics and assess the utility of new biomarkers. We present a novel technique using the concept of Statistical Path Analysis to provide a quantitative framework of cause-effect relationships among the various metabolites involved in the pathway of catecholamine synthesis. This has been used to validate clinical findings in the case of Dopamine-producing paragangliomas. Path analysis also correctly predicts the clinical presentation of patients (normotensive due to the low dominance of Epinephrine sub-pathway). Statistical path analysis, therefore, can serve as a useful tool to quantitatively validate clinical hypothesis on patient cohorts and can serve as a useful aid in identifying new biomarkers.**

Keywords – Statistical Path Analysis; Quantitative Hypothesis Validation; Biomarker Assessment; Dopamine-Predominant Paragangliomas; Methoxytyramine Utility

I. INTRODUCTION

Biochemical tests are often the first means of detecting any abnormality in patients; hence they are extremely important with respect to screening and diagnosis. The basis of biochemical testing is that the secretions of the diseased tissues are different from those of the normal tissue. Concentrations of these secretions, which ideally need to be measured at the specific tumor site, are most conveniently estimated by their concentrations in the blood plasma and urine. The tests are designed based on the biochemical pathway relevant to the disease and their success depends on the sensitivity and specificity of these secretions, which serve as biomarkers for the disease. The changes in concentration are useful for the differentiation between normal and abnormal tissue. The distinction is made mostly based on heuristics and in case of conflicting evidence, by the expert judgment of the clinician. Since the experience and insight of clinicians differ, it is important to have a system that can aid clinicians in their assessment. Biochemical tests are also used in clinical trials where a study group is subject to a treatment and the progress of the study group is compared with an untreated control group. Plasma and urine concentrations of specific metabolites are commonly used to measure the progress of the study group. Sometimes, the change in values gives mixed results, which renders the test inconclusive and the clinical trials of the drug are forced to be abandoned. In such cases, it would be useful to determine the causes of the change (or lack of change) in the metabolite levels which can provide further insight into the action of the drug. In some cases, the drug may be effective but the metabolite levels are inconclusive due to other reasons. In such cases, it would also be useful to identify if such effects are present. Thus, it is necessary to have a framework by which clinicians can make decisions regarding causes of observed conditions, quantitatively validate their heuristics and assess the utility of new biomarkers.

One of the main limitations of clinical studies is the fact that metabolite levels are compared with reference levels, but their role in the biochemical pathway is not used explicitly for diagnoses. Biochemical pathways such as metabolic, regulatory or signal transduction pathways can be viewed as interconnected processes forming an intricate network of functional and physical interactions between molecular species in the cell. The amount of information available on such pathways for different organisms is increasing rapidly. This offers the possibility of performing various analyses on the structure of the full network of pathways for one organism as well as across different organisms, and has therefore generated interest in analysis of pathways. Pathways can also be analyzed to study the metabolic changes that take place when a tumor forms, and can hence be used to predict or diagnose a tumor.

One such important class of tumors are Dopamine-producing paragangliomas which originate in the ganglia of the sympathetic nervous system. These tumors are rare, but are important to diagnose early because of their disposition to produce distant metastases [1, 2]. Patients who have predominantly Dopamine producing tumors are generally normotensive, which poses a significant diagnostic challenge. Further, the sole reliance on the measurement of plasma concentrations of Epinephrine and Norepinephrine or their O-methylated metabolites fails to detect Dopamine predominant tumors. In a recent clinical study [3], it was proposed that Methoxytyramine could be a better marker for Dopamine-predominant paragangliomas. We provide a novel

quantitative framework based on statistical path analysis, which can help quantify such clinical hypotheses and provide further insights in case of conflicting evidence. The experimental background and significance is presented in Section 2. The theory of statistical path analysis as applied to the diagnosis of paragangliomas is presented in section 3. Results and Conclusions are presented in sections 4 and 5 respectively.

II. Analysis Of Biochemical Pathways

The concentrations of catecholamine secretions follow a certain well-understood biochemical pathway, in which each product is formed from its precursor by the action of an enzyme. The biochemical pathway of Dopamine production in the human glands is shown in Fig. 1. In the figure given below, the enzymes governing the various reactions are TH (Tyrosine hydroxylase), AADC (Aromatic amino acid decarboxylase),COMT (Catechol O-methyltransferase), and DBH (Dopamine β-hydroxylase). These tissue metabolites may either be secreted into the bloodstream to form plasma metabolites or eliminated via the kidney to form urinary metabolites. Normally, both plasma and urine concentrations are measured to get the effective concentrations. In a clinical study on the diagnosis of these paragangliomas [3], 120 patients with catecholamine producing tumors were chosen from a larger group of 279 patients with tumors, based on results obtained from biochemical testing. These tests included measurement of plasma concentrations of the catecholamines and their respective O-methylated products. It also included the measurement of plasma DOPA concentration and 24-hour urinary outputs of catecholamines, using the technique of HPLC [12].

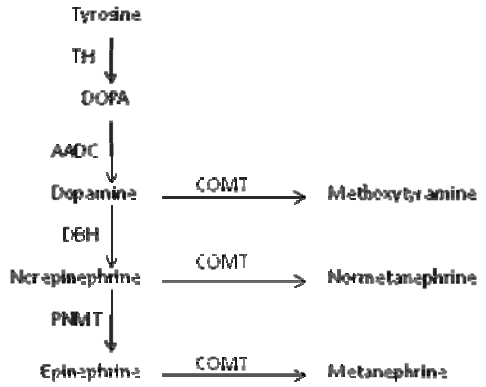

Figure 1. Biochemical Pathway of Dopmaine Synthesis

Through similar studies, reference ranges for the compounds were determined from a combined group of 175 normotensive and 110 hypertensive volunteers. These were used to isolate patients with Dopamine-producing paragangliomas and assess the suitability of the current biomarkers. From the measurements, it was hypothesized that Methoxytyramine could be a more sensitive marker than the current markers, Norepinephrine and Epinephrine. It was also predicted that urinary Dopamine may not be a reliable indicator of Dopamine secretions since it is more likely to result from the renal elimination of plasma DOPA rather than the elimination of plasma Dopamine.

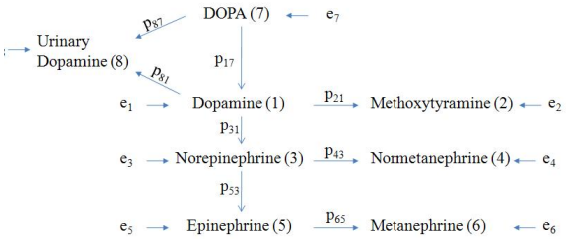

Figure 2. Path Analysis applied to the Dopamine Pathway.

III. Path Analysis

In this work, the above mentioned measurements were analyzed using the method of Statistical Path Analysis. Statistical Path Analysis [6] was developed as a method of decomposing correlations into different components for interpretation of effects and is closely related to multiple regression. This technique allows us to test theoretical propositions about cause-effect relationships without manipulating variables. The path analysis model applied to the Dopamine synthesis pathway is given in Fig. 2.

where

- p_{ij} denotes the path coefficient between precursor 'j' and product 'i'
- e_i denotes the error term which accounts for causes outside the model.

If Z_i denotes the normalized concentration of the i^{th} metabolite in the pathway, the system of equations describing the model can be written as

$$\begin{aligned} Z_7 &= e_7 & Z_3 &= p_{31}Z_1 + e_3 & Z_6 &= p_{65}Z_5 + e_6 \\ Z_1 &= p_{17}Z_7 + e_1 & Z_4 &= p_{43}Z_3 + e_4 & Z_8 &= p_{87}Z_7 + p_{81}Z_1 + e_8 \\ Z_2 &= p_{21}Z_1 + e_2 & Z_5 &= p_{53}Z_3 + e_5 \end{aligned} \quad (1)$$

Computation of path coefficients may be done by determining the correlation coefficients and decomposing them into the following components, each denoting a specific type of effect.

1. Direct effect (DE): the effect of one variable on another, which is represented in the model by a single causal path

2. Indirect effect (IE): a path from one variable to another, which passes through some other variable (the intervening variable or mediator)
3. An unanalyzed component (U): due to lack of information on the direction of causation for a path or due to correlated exogenous variables.
4. Spurious component (S): correlation between two variables due to a common cause.

Therefore, r = DE + IE + U + S

In order to compute the path coefficients for the Dopamine pathway, the measured concentrations of metabolites above the Upper Reference Level (URL) were considered as raw variables. Standardization of raw data was carried out through the calculation of Z scores, using the formula

$$Z = (x - \mu)/\sigma \quad (2)$$

Where μ is the mean of the distribution, σ the standard deviation and x the raw value.

For every precursor (a)-product (b) metabolite pair within the biosynthetic pathway, a correlation coefficient was computed as follows:

$$r_{ab} = (1/N) \sum Z_a Z_b \quad (3)$$

Substituting (1) in (3) and applying the following simplifications

- $\sum Z_i e_j = 0$ since there is no effect of e_i on Z_j(for $i \neq j$)
- $(1/N) \sum Z_i Z_i = 1$ since $Z_i$are normalized,

We obtain a system of equations relating path coefficients and the correlation coefficients.

$$r_{71} = p_{17} \qquad r_{12} = p_{21} \qquad r_{13} = p_{31}$$
$$r_{34} = p_{43} \qquad r_{35} = p_{53} \qquad r_{56} = p_{65}$$
$$r_{32} = p_{31}p_{21} \qquad r_{54} = p_{53}p_{43} \qquad r_{78} = p_{87} + p_{81}p_{17}$$
$$r_{18} = p_{87}\, p_{17} + p_{81} \quad (4)$$

An alternate method of arriving at (4) is by decomposing the correlations into various effects with respect to the model (Fig. 2) and is carried out as follows:

- The correlation between DOPA and DA, DA and MTY, DA and NE, NE and NMN, NE and EPI, EPI and MN can be decomposed into their respective direct effects. For example, $r_{71} = p_{17}$
- The correlation between DOPA and urinary DA, DOPA and the O-methylated metabolites, DA and NMN/MN, NE and MN, can be expressed in terms of their respective direct and indirect effects. For example, the indirect component of r_{72} is $p_{17} \times p_{21}$
- The correlation between NE and MTY (sharing a common cause DA), that between EPI and NMN (sharing a common cause NE) and that between plasma DA and urinary DA (sharing a common cause plasma DOPA), can be decomposed into their direct, indirect and spurious effects. For example, the spurious component of r_{32} is $p_{21}xp_{31}$

Thus knowing the values of correlation coefficients, the values of the respective path coefficients can be computed.

IV. Results

A. Calculation of path coefficients

The method of statistical path analysis was applied to data obtained from 9 patients with a confirmed diagnosis of Dopamine-predominant paraganglioma. The plasma and urinary concentrations of metabolites involved in the Dopamine pathway were measured by the method of HPLC. The plasma concentrations of the values of the metabolites measured are given in Table 1. The Upper Reference Level (URL) for the various compounds is also given. All concentrations reported are in nanograms/litre. It can be seen that all patients showed greater elevations in the plasma concentrations of Dopamine and Methoxytyramine as compared to the other compounds in the pathway. It can also be seen that the concentration of plasma DOPA in these patients is very high, which led to the suspicion that the contribution of the plasma DOPA is much higher than that of plasma Dopamine to urinary Dopamine [3].

TABLE I. Plasma Concentrations of Metabolites

Patients	Plasma concentrations						
	MTY	***NMN***	***MN***	***DA***	***NE***	***EPI***	***DOPA***
1	93	60	15	160	571	16	1805
2	419	29	38	600	118	44	1490
3	488	241	38	776	717	17	1615
4	634	624	44	2092	2784	12	2784
5	1189	1273	77	2669	2229	29	7988
6	2374	1053	26	3705	5131	57	1033
7	7022	4636	42	17749	14935	106	18221
8	9956	192	2	27942	1030	15	4782
9	12003	1259	5	55619	6553	177	48156
URL	14	112	61	58	496	83	2529

From these values, the correlation coefficients for all metabolite pairs which have an immediate precursor-product relationship were determined (Table 2). The path coefficients corresponding to these immediate pairs were

obtained by solving the system of equations (4). The path coefficients so obtained are given in Figure 3.

B. Model Preditions and Analysis

From the quantitative results, we observe the following:

1) The correlation between plasma Dopamine and Methoxytyramine (r=0.8452), plasma Dopamine and DOPA (r=0.7976), and plasma Norepinephrine and Normetanephrine (r=0.8586) show a strongly positive association, indicating that an increase in one of the compounds strongly contributes to an increase in the other.

2) The correlation between plasma Dopamine and Norepinephrine (r=0.3358), plasma Norepinephrine and Epinephrine (r=0.5827), urinary Dopamine and plasma DOPA (r=0.599) and urinary Dopamine and plasma Dopamine (r=0.3113) show a weak positive association. The increase in one of the compounds causes an increase in the other, but to a lesser extent than in the first case.

3) The correlation between plasma Epinephrine and Metanephrine (r=-0.2474) indicates little or no association. Increase or decrease in one compound maybe associated with both an increase and a decrease in the other compound.

C. Inferences

Since the path coefficients indicate a cause-effect relationship, the following can be deduced:

1) The dominant pathway corresponds to DOPA→Dopamine→Methoxytyramine. Hence, Methoxytyramine can serve as the most specific marker for patients of this class of tumors.

2) The path coefficient for the production of methoxytyramine from dopamine (0.8452) is greater than the path coefficient for the production of norepinephrine from dopamine (0.3358). This indicates that the dopamine contributes more to the production of methoxytyramine than to norepinephrine. Moreover since most of the methoxytyramine comes from dopamine (r=0.8452), plasma free methoxytyramine serves as a sufficiently good marker for the detection of predominantly dopamine producing paragangliomas. This also verified the results obtained by biochemical testing carried out by Graeme Eisenhofer et al [3].

3) The path coefficients can also be used to compute the primary cause of increase in urinary Dopamine. To do this, we need to consider two sub-pathways: plasma DOPA → urinary Dopamine and plasma Dopamine → urinary Dopamine. The effect of plasma DOPA on urinary Dopamine is computed by considering both the direct and indirect effects of plasma DOPA on urinary DA. Total Effect = Direct Effect + Indirect Effect = 0.599 + (0.7976)x(0.3113) = 0.8473. The effect of plasma Dopamine on urinary Dopamine is computed by considering direct effect (p32) and spurious effect (given by p21 x p31). Hence, the total effect of plasma Dopamine on urinary Dopamine is 0.3113 + 0.7976 x 0.599 = 0.7891. This also represents the correlation between plasma Dopamine and urinary Dopamine. Thus, it can be seen that Plasma DOPA is a more reliable indicator of urinary Dopamine than plasma Dopamine. This has also been validated by expert insights.

4) The path coefficients of the Norepinephrine and Epinephrine sub-pathways have very low dominance (low path coefficients). This indicates that in predominantly Dopamine producing paragangliomas, the possibility of Norepinephrine being formed from the Dopamine that is produced by the tumor cells is low (r=0.3358). Since Norepinephrine is primarily responsible for the production of Noradrenaline and Adrenaline, the levels of Norepinephrine and Epinephrine are not sufficiently high to result in hypertension, in most of the cases. This justifies the normotensive nature of most of the patients having predominantly Dopamine producing paragangliomas. This normotensive nature poses a diagnostic challenge.

5) The low dominance of the Norepinephrine pathway also indicates that Norepinephrine and Epinephrine are not suitable markers for Dopamine-predominant paragangliomas. Since current tests mainly rely on the levels of Norepinephrine and Epinephrine to carry out diagnosis, this also explains the reason behind the large number of cases of Dopamine-producing paragangliomas which were missed during clinical tests.

Sensitivity of Methoxytyramine as a marker for dopamine-predominant paragangliomas is evident by the the observed 104-fold increase [3]. By combining it with a quantitative measure of specificity (via path analysis), it is possible to establish quantitatively the utility of Methoxytyramine in diagnosing dopamine-producing paragangliomas.

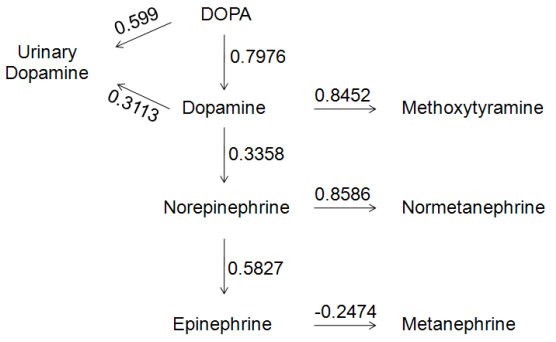

TABLE II. CALCULATED CORRELATION COEFFICIENTS

r_{12}	r_{13}	r_{34}	r_{35}	r_{54}	r_{56}	r_{17}	r_{32}	r_{78}	r_{18}
0.8452	0.3358	0.8586	0.5827	0.46697	-0.2474	0.7976	0.4156	0.8473	0.78906

V. CONCLUSION

We have presented a novel framework for quantitative validation of clinical studies. The framework based on statistical path analysis has been found to validate clinical findings on a cohort of patients. The predictions of the model have been found to agree very well with clinical findings in terms of predicting the utility of Methoxytyramine as a useful marker, the insensitivity of urinary Dopamine as a marker and also additionally the clinical symptoms. This promising method can be used further for verification of biochemical diagnosis of diseases on patient cohorts.

It can also be used to compare the utility of various biomarkers quantitatively for a patient group with a known diagnosis. Path analysis can give an accurate indication of the specificity of biomarkers. Currently, there is no quantitative measure of specificity and path analysis may be able to serve as one. However, the sensitivity of biomarkers cannot be predicted by statistical path analysis yet and may need additional measures.

One of the main challenges faced by clinicians is in performing clinical trials. Comparisons between control and test groups are often done by comparing the mean and standard deviation of parameters. By employing statistical path analysis models, it may be possible to judge not only the improvements in the study group, but also whether the biochemical changes justify such an improvement.

Thus, statistical path analysis is a powerful technique which can be extended for multiple uses benefiting clinical diagnosis and measurement of therapy response.

REFERENCES

[1] MedicineNet.com http://www.medicinenet.com/pheochromocytoma/article.htm

[2] AshleyB.Grossman, MERCK manuals, Last full review/revision November 2007 http://www.merck.com/mmpe/sec12/ch153/ch153h.html

[3] Graeme Eisenhofer, David S. Goldstein, Patricia Sullivan, Gyorgy Csako, Frederieke M. Brouwers, Edwin W. Lai, Karen T. Adams and Karel Pacak, "Biochemical and Clinical Manifestations of Dopamine-Producing Paragangliomas: Utility of Plasma Methoxytyramine", Journal of Clinical Endocrinology & Metabolism, Vol. 90, No. 4 2068-2075

[4] Graeme Eisenhofer, Jacques W.M. Lenders, David S. Goldstein, Massimo Mannelli, Gyorgy Csako, McClellan M. Walther, Frederieke M. Brouwers and Karel Pacak, "Pheochromocytoma Catecholamine Phenotypes and Prediction of Tumor Size and Location by Use of Plasma Free Metanephrines", Clinical Chemistry 51: 735-744, 2005. First published February 17, 2005; 10.1373/clinchem.2004.045484

[5] Havekes B, Lai EW, Corssmit EP, Romijn JA, Timmers HJ, Pacak K, "Detection and treatment of pheochromocytomas and paragangliomas: current standing of MIBG scintigraphy and future role of PET imaging", Q J Nucl Med Mol Imaging. 2008 Dec;52(4):419-29.-PubMed

[6] Path analysis http://luna.cas.usf.edu/~mbrannic/files/regression/Pathan.html

[7] Research Methods and Statistics PESS202 Lecture and Commentary Notes-2000 , Chapter 4: Analysing the Data, Part II : Descriptive Statistics http://www.une.edu.au/WebStat/unit_materials/c4_descriptive_statistics/z_scores.htm

[8] Path analysis http://www.gseis.ucla.edu/courses/ed230b/notes/handout7.pdf

[9] Correlation:Interpretations http://www.visualstatistics.net/Visual%20Statistics%20Multimedia/correlation_interpretation.html

[10] Steve Simon, 2005-08-18 http://www.cmh.edu/stats/definitions/correlation.htm

[11] JW Lenders, G Eisenhofer, I Armando, HR Keiser, DS Goldstein and IJ Kopin, "Determination of Metanephrines in plasma by liquid chromatography with electrochemical detection",Clinical Chemistry, Vol 39, 97-103, Copyright © 1993 by American Association for Clinical Chemistry.

[12] Path analysis http://people.eku.edu/falkenbergs/psy862/pathdefs.htm

[13] Moses E. Olobatuyi, "A User's guide to path analysis", Chapter seven (Page 133-149)