Within-group and between-group correlation: Illustration on noninvasive estimation of intracranial pressure

Caren Marzban, Paul R. Illian, David Morison, Pierre D. Mourad

***Abstract*—The linear association between two continuous quantities is often assessed in terms of Pearson's correlation coefficient, r. However, if the data are not homogeneous, i.e., consist of groups, then it is important to decompose the "total" correlation into components that measure the correlation within the groups, and the correlation between the groups. Although this type of decomposition is relatively well-known, it appears to be rarely practiced. To illustrate the importance of distinguishing between the three notions of correlation, we compute the correlations between Intracranial Pressure (ICP) and three quantities: 1) maximum Flow Velocity, FV (obtained from Transcranial Doppler), 2) Arterial Blood Pressure (ABP), and 3) Zero-Flow Pressure (ZFP). We show that FV is useless for predicting mean (over time) ICP for patients, but is more useful for monitoring ICP across time for a given patient. Similarly for ABP. By contrast, ZFP is useful for predicting mean ICP for different patients, but useless for monitoring an individual's ICP across time.**

***Index Terms*—Performance evaluation, correlation,, noninvasive treatment, intracranial pressure sensors**

INTRODUCTION

GIVEN data on two continuous quantities x and y, the strength of the (linear) association between them is often assessed by a scatterplot of y vs. x. A strong linear pattern along the diagonal of the scatterplot indicates strong association. If x and y are observations and corresponding predictions, respectively, then the overall position of the points on a scatterplot, relative to the diagonal, is a measure of the accuracy of the predictions, and the amount of scatter about the diagonal gauges the precision of the predictions. A popular, scalar measure of the latter is Pearson's correlation coefficient [1], here denoted r.

If the data are not homogeneous (i.e., if the scatterplot scatterplot displays clusters/groups), then there are three different correlations: 1) correlation, ignoring the existence of groups altogether; 2) correlation within the groups; and 3) correlation across the groups. Borrowing from the language of analysis of variance [2], in this paper we refer to these as *total correlation*, *within-group correlation*, and *between-group correlation*, respectively. Each of these three correlations has a different meaning, and the appropriate choice depends on the problem at hand. For instance, a large within-group correlation between x and y suggests that one can predict y from x, but not necessarily across the groups. The manner in which knowledge of x in one group affects y in another group is related to the between-group correlation.

The "causes" of groups in scatterplots can vary, but a common situation in which groups occur is when multiple measurements (and predictions) are made on multiple subjects, in which case each group corresponds to a unique subject.

The notion of within-group and between-group correlation is a natural extension of the same notion pertaining to variance [2]. Both are relatively well-known, and even common-knowledge in many circles. However, as shown in the next section, many practitioners often fail to acknowledge the distinction, and so, either compute the total correlation even in the presence of groups in the data, or compute only one of the components. In all of these instances valuable information is lost.

One way to illustrate the importance of distinguishing between the total and within-group correlation is through Simpson's paradox [3]: Consider the simulated data shown in Fig 1. The r for each of the two groups (i.e., measuring correlation within-group) is 0.9. But the total r for the entire data set is zero. In other words, by examining only the total correlation, one can miss the fact that the correlation within group is very high. More realistic examples are provided below.

I. INTRACRANIAL PRESSURE (ICP)

ICP is an important factor in monitoring patients who have experienced head trauma as well as other conditions [4], [5].

Manuscript received Jan. 20, 2013. This work has received support from National Institutes of Health (Grant R43NS46824-01A1); National Space Biomedical Research Institute (Grant SMS00701-2009-513); and PhysioSonics Incorporated. P. D. Mourad has a financial interest in PhysioSonics.

C. Marzban is with the Applied Physics Laboratory (APL), and the Department of Statistics, University of Washington, Seattle, WA 98195-4322 USA. Phone: 206.221.4361; fax: 206-543-1301; e-mail: marzban@stat.washington.edu). Corresponding author.

P. R. Illian and D. Morison work at the APL, University of Washington, Seattle, WA 98195 USA. Emails: rillian@apl.washington.edu and davidm@apl.washington.edu

P. D. Mourad is with the APL, Department of Neurological Surgery, and Department of Bioengineering, at the University of Washington, Seattle, WA 98195. Email: pierre@apl.washington.edu

Direct measures of ICP are invasive, requiring the placement of pressure sensors within the cranium. Therefore, there is great interest in finding noninvasive measures of ICP [6]-[28]. Three predictors of ICP proposed in the literature are maximum Flow Velocity (FV) as measured by Transcranial Doppler, Arterial Blood Pressure (ABP), and Zero-Flow Pressure (ZFP) [6,7,11,26]. It is important to point out that none of these predictors are clinically useful, especially when used one at a time. However, they suffice to illustrate the importance of separately considering within-group and between-group correlations.

As shown below, these data are not homogeneous, and so, there exist different notions of correlation, each with a different meaning and utility depending on the problem at hand. If one is interested in predicting the mean (over time) ICP for a patient, then the between-group correlation is appropriate for assessing the quality of the predictions. On the other hand, if one is interested in predicting ICP, for a given patient, across time, then the appropriate quantity is the correlation within-group.

The ICP literature abounds with instances where correlation between ICP and some other quantity is examined. The quantity may be a noninvasive estimate of ICP (e.g., ZFP), or simply a predictor of ICP (e.g., FV or ABP). In the former, the correlation is employed to assess the quality of the predictions, while in the latter it is used for determining the predictive strength of the predictor. All of the aforementioned ICP-related articles [6]-[28] report some value of r, but they report either the total correlation, or only one of between-group or within-group correlation. Some [11,17,24,27] of the aforementioned works even report scatterplots which clearly display groups, and yet they still compute only the total correlation. As such, correlation is not being properly interpreted.

The next section demonstrates the consequences of this type of misinterpretation in a concrete example. Other illustrations of the notions of between-group and within-group correlation can be found in [28-30].

II. Data and Method

We collected data on FV, ABP, ICP and ZFP for 104 patients. Details of the data are presented in [26]. In order to examine the association between x=FV and y=ICP, and between x=ABP and y=ICP, 200 random samples (at different times) were taken from each patient's data. Most measures of ZFP involve linearly extrapolating ABP to the point where FV is zero (hence, the name zero-flow-pressure). To study the association between x=ZFP and y=ICP, 200 random samples (at different times), across a time window 7.5 seconds long, were taken from each patient's data, and compared with the mean of ICP across the same time window. Said differently, the study of correlations here involves 104 groups, each consisting of 200 points.

The total correlation was measured by computing r across the entire data (i.e., 104*200 cases), ignoring the groups altogether. Computing r within each of the groups (i.e., across 200 cases), and then averaging them across the 104 patients was taken as the measure of within-group correlation. The histogram of the within-group correlations is also examined, because there are instances in which this histogram is skewed, and so, the mean of the within-group correlations is not necessarily the best summary measure of the distribution. The between-group correlation is obtained by averaging the x and y values for each group across the 200 points (leading to a scatterplot where each group is replaced by the average of the corresponding x and y), and then computing r across the 104 patients.

III. Results

Table 1 shows all of the correlations. (All but two are associated with very low p-values, suggesting that the true/population correlations are nonzero. The two correlations whose zero value cannot be ruled out, because of high p-values, are very low, and marked with a *.)

If one were to consider only the total correlation, then the conclusion would be that FV is not correlated with ICP, but ABP and ZFP are correlated with ICP. This would be an incorrect conclusion; the true nature of the correlation is more complex; see the top panels of Fig.2 and Fig. 3. For instance, in spite of being apparently uncorrelated (according to the total r), the value of the within-group correlation suggests that FV is correlated with ICP, for individual patients. In other words, monitoring FV for a given patient provides a noninvasive estimate of ICP for that patient. The low value of the between-group correlation implies that one cannot estimate a patient's mean (over time) ICP from knowledge of the mean FV.

ABP offers a somewhat different conclusion: The non-zero value of the total correlation suggests that ABP may be a predictor of ICP, albeit a week predictor because of the relatively low value of the correlation (i.e., 0.22). However, the between-group correlation is about 0.5, and is therefore a much better predictor if one employs ABP for monitoring a given patient's ICP. The between-group correlation is still relatively low (0.21); although it will provide some measure of predictability of mean ICP across patients, the quality of the prediction is likely to be poor.

By contrast with ABP and FV, the pattern of correlations between ZFP and ICP is "reversed;" the total correlation is relatively high (0.51), but all of that is due to between-group correlation (0.52). The within-group correlation is nearly zero. As such, ZFP is a poor predictor of ICP for a given patient's ICP as a function of time, but it is a much better predictor when a patient's mean (across time) ICP is desired.

The scalar numbers in Table 1 can be diagnosed further by examining Fig. 2 and Fig. 3. The top panel in Fig. 2 shows the scatterplot of FV and ICP for all 104 patients (in different colors). The low value of the total correlation noted in Table 1 is substantiated in this scatterplot. The histogram of the correlation coefficient between FV and ICP for each patient is shown in the middle panel of Fig. 2. It is the mean of this histogram, which is reported in Table 1 as the within-group

correlation. Given the skewed nature of this histogram, it is evident that the mean under-estimates the correlation within-group. For example, the median of the within-group correlations is 0.46 (as compared to the mean which is 0.43). The scatterplot of the patient means (bottom panel) confirms the low value of the between-group correlation reported in Table 1.

The analogous figures for examining the correlations between ABP and ICP are similar to those in Fig. 2, and are therefore not shown.

Fig. 3 shows the analogous figures for examining the correlations between ZFP and ICP. The top panel shows the scatterplot of ICP vs. ZFP for all 104 patients. The distribution of within-group correlations (middle panel) shows that the within-group correlations span the full range from nearly -1 to nearly +1. In other words, for some patients, ZFP is a nearly perfect predictor of ICP, and for some patients it is not. Barring any information that can a priori discriminate between these classes of patients, on the average the within-group correlation is zero, and so, ZFP is not useful for predicting ICP across time for a given patient. The bottom panel shows the scatterplot of the group-means of ZFP and ICP, and confirms the relatively large between-group correlation shown in Table 1.

IV. Conclusion and Discussion

The conclusions of this study are two-fold: 1) We have demonstrated that examining the total correlation coefficient can lead to incorrect or misleading conclusions regarding the nature of the underlying relationship between two variables. It is important to consider the within-group and between-group correlations as well, because they have different practical implications. 2) By examining a large data set involving the FV, ABP, ZFP, and ICP of 104 patients, we have shown that there exists a complex relationship regarding the strength of these quantities in terms of estimating ICP; briefly, FV and ABP are good predictors of ICP within-patient (i.e., for monitoring the ICP of a given patient across time), while ZFP is a good predictor of ICP between-patients.

In [26] we proposed several revisions to the methodology for computing ZFP all of which were shown to improve the quality of the ICP prediction. Several measures of quality were employed, but none were decomposed in the fashion proposed here. We have now performed that decomposition, but the conclusions in [26] are unaffected, and so the results are not discussed here.

The within-group and between-group correlations computed here may be considered naive, because they are not based on a rigorous decomposition of covariance into within-group and between-group terms. We have performed that decomposition, and found that the results are virtually the same as those reported here. A technical memo, describing that approach is available upon request from the corresponding author.

It is worth mentioning that in this paper, all references to "good" or "bad" predictors of ICP are in the relative sense. In fact, none of these predictors are sufficiently good to be of clinical use. The quality of ZFP (the best of the three predictors considered here) as an ICP predictor has been thoroughly examined in [26].

References

[1] N. R. Draper and H. Smith, *Applied Regression Analysis*, 3rd ed. New York: John Wiley & Sons, 1998.

[2] D. C. Montgomery, *Design and Analysis of Experiments*, 6th ed. New York: John Wiley & Sons, 2005.

[3] C. H. Wagner, "Simpson's paradox in real life." *The American Statistician*, Vol. 36, pp. 46-48, 1982.

[4] R. Bullock , et al., "Guidelines for the management of severe head injury. Brain Trauma Foundation," *J Neurotrauma*, Vol. 17, pp. 451-553, 2000.

[5] H. C. Patel ,O. Bouamra , M. Woodford , A. T. King, D. W. Yates, F. E. Lecky, "Trauma Audit and Research Network. Trends in head injury outcome from 1989 to 2003 and the effect of neurosurgical care: an observational study," *Lancet* , Vol. 366, 2005, pp. 1538-1544.

[6] R. Burattini, P. Borgdorff, D. R. Gross, B. Baiocco, and N. Westerhof, "Systematic autoregulation counteracts the carotid baroreflex," *IEEE Trans. Biomedical Engineering,* Vol. 38, pp. 48-56, 1991.

[7] A. Weyland, et al., "Cerebrovascular tone rather than intracranial pressure determines the effective downstream pressure of the cerebral circulation in the absence of intracranial hypertension." *J Neurosurg Anesthesiol ,* Vol. 12, pp. 210–6, 2000.

[8] C. Thees , et al, "Relationship between intracranial pressure and critical closing pressure in patients with neurotrauma," *Anesthesiology*, Vol. 96**,** pp. 595–599, 2002.

[9] M. Aboy, et al, "Significance of intracranial pressure pulse morphology in pediatric traumatic brain injury," in *Proc. 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, Cancun, Mexico, September 17-21,2003, pp. 2491-2494.

[10] B. Schmidt , M. Czosnyka , A. Raabe , H. Hilal , J. J. Schwarze, D. Sackerer, D. Sander, J. Klingelhöfer , "Adaptive Noninvasive Assessment of Intracranial Pressure and Cerebral Autoregulation," *Stroke,* Vol. 34, pp. 84-89, 2003.

[11] W. Buhre, F. R. Heinzel, S. Grund, H. Sonntag, A. Weyland , "Extrapolation to zero-flow pressure in cerebral arteries to estimate intracranial pressure." *Br J Anaesth .,* Vol. 90, pp. 291-295, 2003.

[12] R. Hornero, M. Aboy, D. Abásolo, J. McNames, and B. Goldstein, "Interpretation of approximate entropy: Analysis of intracranial pressure approximate entropy during acute intracranial hypertension," *IEEE Trans. Biomedical Engineering,* Vol. 52, pp. 1671-1680, 2005.

[13] A. Ragauskas, G. Daubaris, A. Dziugys, V. Azelis, V. Gedrimas , "Innovative non-invasive method for absolute intracranial pressure measurement without calibration," *Acta Neurochir Suppl.*, Vol. 95, pp. 357-361, 2005

[14] M. Czosnyka, P. Smielewski , I. Timofeev, A. Lavinio,E. Guazzo, P. Hutchinson , J. D. Pickard, " Intracranial Pressure: More than a number," *Neurosurg Focu*s, Vol. 22, pp. 1-7, 2007.

[15] Behrens A, Lenfeldt N, Ambarki K, Malm J, Eklund A, Koshinen L. Transcranial Doppler pulsatility index: Not an accurate method to assess intracranial pressure. *Neurosurgery* 2010; **66:** 1050-1057.

[16] P. Xu, M. Kasprowicz, M. Bergsneider, and X. Hu, " Improved Noninvasive Intracranial Pressure Assessment With Nonlinear Kernel Regression," *IEEE Trans. Information Technology in Biomedicine*, Vol. 14, pp. 971-978, 2010.

[17] M. Chacón, C. Pardo, C. Puppo, M. Curilem, and J. Landerretche, "Noninvasive intracranial pressure estimation using support vector machine," *32nd Annual International Conference of the IEEE Bioengineering in Medicine and Biology Society (EMBS)*, Buenos Aires, Argentina, August 31 - September 4, 2010, pp. 996-999.

[18] W. Chen , C. Cockrell , K. R. Ward , K. Najarian, "Intracranial pressure level prediction in traumatic brain injury by extracting features from multiple sources and using machine learning methods," *IEEE*

International Conference on Bioinformatics and Biomedicine, Hong Kong, Dec.18 – 21,2010, B510.

[19] J. McNames, et al, "Precursors to rapid elevation in intracranial pressure," *Proceedings of the 23rd Annual EMBS International Conference,* Istanbul, Turkey, Oct. 25-28, 2010, pp. 3977-3980.

[20] L. Qiu, L. Xu, Y. Wang, "Modeling of the interaction between intracranial pressure and cerebral blood flow," *3rd International Conference on Biomedical Engineering and Informatics,* Yantai, China, Oct 16-18, 2010, pp.1217-1220, 2010.

[21] P. W. Angriyasa, Z. Rustam, and W. Sadewo, "Non-invasive intracranial pressure classification using strong jumping emerging patterns," in International Conference on Advanced Computer Science and Information Systems (ICACSIS), Jakarta, Indonesia, December17-18, 2011, pp.377-380.

[22] M. C. Aoi, B. J. Matzuka, and M. S. Olufsen, "Toward online, noninvasive, nonlinear assessment of cerebral autoregulation," 33rd Annual International Conference of the IEEE Engineering *in Medicine and Biology Society (EMBS)* , Boston, Massachusetts USA, August 30 - September 3, 2011, pp. 2410-2413.

[23] An. Calisto, M. Galeano, S. Serrano, Am. Calisto ,and B. Azzerboni, "A new approach for investigating intracranial pressure signals: filtering and morphological features extraction from continuous recording, *IEEE Trans. Biomedical Engineering (TBME)*, Vol. 99, pp. 1-7, 2012.

[24] Kashif FM, Verghese GC, Novak V, Czosnyka M, Heldt T. Model-based noninvasive estimation of intracranial pressure from cerebral blood flow velocity and arterial pressure. *Sci Transl Med* 2012; **4:** 1-9

[25] S. Kim, et al, "Noninvasive intracranial pressure assessment based on a data-mining approach using a nonlinear mapping function," *IEEE Trans. on Biomedical Engineering,* Vol. 59, pp. 619-626, 2012.

[26] C. Marzban, P. R. Illian, D. Morison, A. Moore, M. Kliot, M. Czosnyka, P. Mourad, "A method for estimating zero-flow pressure and intracranial pressure," *J Neurosurg*

[27] Yang S, Kalpakis K, Mackenzie CF, Stansbury LG, Stein DM, Scalea TM, Hu PF. Online recovery of missing values in vital signs data streams using low-rank matrix completion. 11th International Conference on Machine Learning and Applications ICMLA 2012, Dec.12-15, Boca Raton, FL.

[28] W. Ting and Z.Shiqiang, "Study on linear correlation coefficient and nonlinear correlation coefficient in mathematical statistics," *Studies in Mathematical Sciences*, Vol. 3, pp. 58-63, 2011.

[29] J. M. Bland and D. G. Altman, "Calculating correlation coefficient with repeated observations: Part I, correlation within subjects," Statistics Notes in *The British Medical Journal*, Vol. 310, pp. 446, 1995.

[30] V. Heuchel, "Between and within subject correlations and variances for certain semen characteristics in fertile men," *Andrologia,*Vol .15, pp. 171-176, 1983.

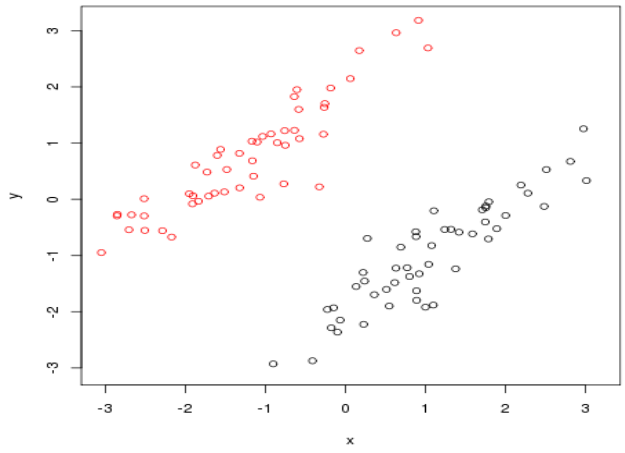

Fig. 1. Illustration of Simpson's paradox. The correlation coefficient between *x* and *y* within each cluster is 0.9, but the overall correlation coefficient is 0.

TABLE I
CORRELATION TOTAL, WITHIN, AND BETWEEN GROUPS

	Total	Within	Between
ICP and FV	0.042	0.43	-0.04*
ICP and ABP	0.22	0.49	0.21
ICP and ZFP	0.51	0.07*	0.52

* non-significant at α=0.05

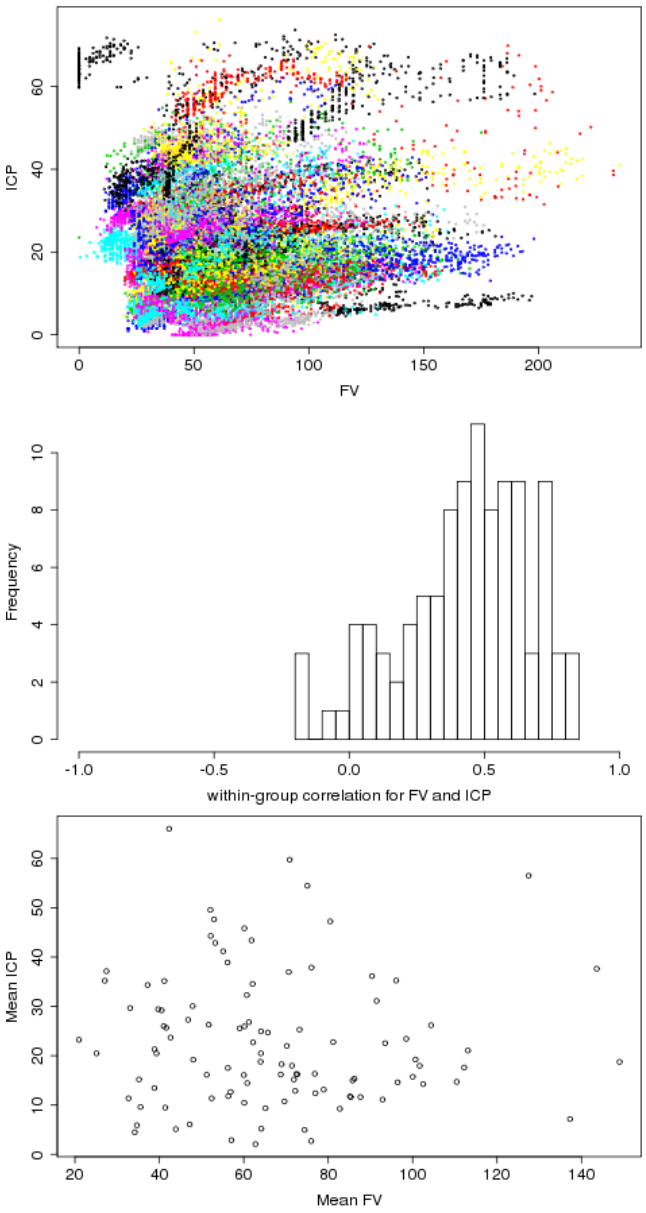

Fig. 2. Top panel: Scatterplot of 200 measurements of ICP and FV for 104 patients (in different colors). Middle Panel: Histogram of the 104 within-group correlations, and scatterplot displaying the between-group correlation.

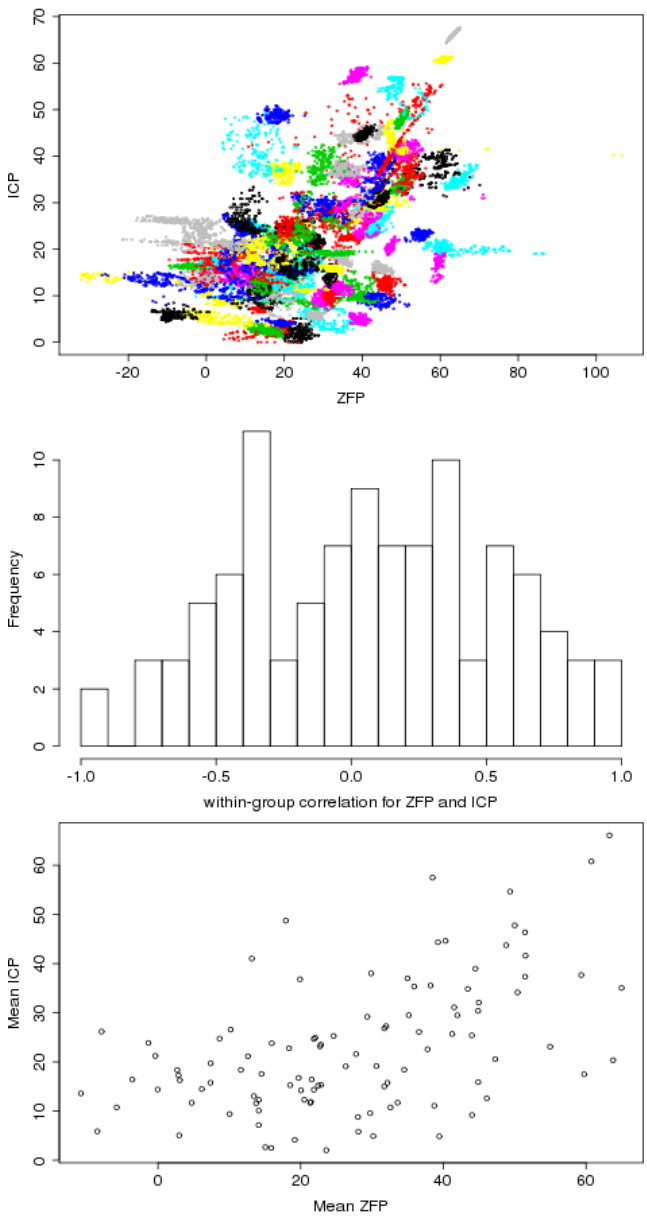

Fig. 3. Same as Fig.2 but for ICP and ZFP.