0008-3194/2007/175–185/$2.00/©JCCA 2007

Reliability of three methods of computer-aided thermal pattern analysis

John Hart, DC, MHS[1]
Bernard Omolo, PhD[2]
W.R. Boone, PhD, DC[3]
Charlene Brown, CT[4]
Andrew Ashton, BS, DC[5]

Background: *The objective of this study was to assess three methods of computer-aided thermal pattern analysis for a) examiner reliability, b) inter-method differences, and c) determine which method yields the highest percent-similarity between paired test-retest scans.*

Methods: *Three examiners compared two sets of thermal scans from the same 30 subjects using three different methods of scan alignment. The results were evaluated by the Intraclass Correlation Coefficient and the Wilcoxon signed-rank test, at the 5% level of significance.*

Results: *Intra and inter-examiner ICC scores for all methods were acceptable (> 0.75). There were no statistically significant differences (at the Bonferroni-corrected level of significance of 0.0004%) in percent similarity of the scans between the three methods*

Conclusions: *The results contribute evidence to the reliability of TPC program software. Manually aligning the readings plays an important role in obtaining precise TPC percent-similarities.*
(JCCA 2007; 51(3):175–185)

KEY WORDS: thermography, vertebral subluxation, reliability

Contexte : *L'objectif de cette étude était d'évaluer trois méthodes d'analyse de profil thermique assistée par ordinateur à des fins de a) fiabilité des examinateurs, b) comparaison entre les méthodes et c) détermination de la méthode produisant le taux le plus élevé de similitude entre les scintigrammes test-retest jumelés.*

Méthodes : *Trois examinateurs ont comparé deux jeux de scintigrammes thermiques des 30 mêmes sujets à l'aide des trois différentes méthodes d'alignement de scintigramme. Les résultats ont été évalués par le coefficient de corrélation intraclasse et le test des rangs signés de Wilcoxon, avec un taux de signification de 5 %.*

Résultats : *Les résultats du coefficient de corrélation intraclasse intra et inter-examinateurs de toutes les méthodes étaient acceptables (> 0,75). Aucune différence statistique significative n'a été relevée (au niveau de signification de 0,0004 % (Bonferroni-corrigé)) en terme de taux de similitude des scintigrammes entre les trois méthodes.*

Conclusions : *Les résultats démontrent la fiabilité du logiciel du programme TPC. Le réglage manuel des lectures joue un rôle important dans l'obtention de similarités en pourcentage de TPC.*
(JACC 2007; 51(3):175–185)

MOTS CLÉS : thermographie, subluxation vertébrale, fiabilité

Contact: John Hart, DC, MHS, Sherman College of Straight Chiropractic, P.O. Box 1452, Spartanburg, South Carolina 29304. Phone: 864-578-8770, ext. 232, Fax: 864-599-4858, jhart@sherman.edu
[1] Assistant Director of Research, Sherman College of Straight Chiropractic, Spartanburg, South Carolina.
[2] Math and Computer Science, University of South Carolina-Upstate, Spartanburg, South Carolina.
[3] Director of Research, Sherman College of Straight Chiropractic, Spartanburg, South Carolina.
[4] Administrative Assistant, Research Department, Sherman College of Straight Chiropractic, Spartanburg, South Carolina.
[5] Private practice.

Introduction

Most chiropractic definitions of vertebral subluxation include a misalignment component and a neurological component.[1] One method of neurological assessment proposed involves thermography. One of the two methods of interpretation of thermography consists of assessing the amount of bilateral temperature asymmetry. This method presumes that the greater the asymmetry the more likely the patient's overall neurology has been compromised.[2] Thermal asymmetry has been noted in a number of health problems, including peripheral nerve involvement.[2, 3] Another method of thermography interpretation involves determining if the asymmetries are persistent; that is, whether the temperature differentials persist as a *pattern*.[4] Although many factors can influence skin temperature, such as ambient room temperature, sunburn, and blemishes, thermal pattern analysis is based on the following general concepts:

1. Skin temperature is largely regulated by the autonomic nervous system.[5]
2. Dynamic skin temperatures reflect a dynamic healthy nervous system evoking adaptive responses to the environment.[4]

Used by chiropractors since the early 1940s,[6] thermal pattern analysis has been applied in case studies[7,8] as well as in a study regarding thermal equilibration.[9] Since thermal pattern analysis is one method used by some chiropractors for assessing the neurological component of the vertebral subluxation, the present investigation was designed to assess intra- and inter-examiner reliability in regard to applying three different methods of aligning scans (graphs) for software analysis.

As with many clinical tests, interpretation can be problematic. Visual inspection of thermal patterns tends to be a subjective process. To render thermal pattern analysis more objective, Stewart et al.[10] developed software concepts for a pattern calculator in 1989. Hart and Boone[11] proposed a non-computer method (using a standard ruler) to demonstrate the basic quantifiable aspects of assessing thermal patterns. Later, Owens and Stein[9] developed thermal pattern calculator (TPC) software which has been used in a recent study on thermal equilibrium.[9] The TPC is designed to analyze thermal scans.[12] As part of this development, Owens, et al.[13] compared several algorithms to ascertain the best method of computer-aided calculations of temperature data. They found the method of Stewart et al., [10] where the slope (r%) of the readings is assessed for percent similarity, to be the most congruent. Percent similarity or *pattern* refers to the amount of slope similarity between two scans being compared. The statistical description for determining percent-similarity in the TPC is presented in the publication by Owens et al.[12] Briefly, the Stewart et al. method[10] uses the Pearson product moment (r) statistic, comparing 10 rows of data points (temperatures) from each of two scans. With eight degrees of freedom, an r value of .632 would be significant at the 0.05 alpha level.[12] Two scans having a percent similarity of 55% would have less amount of slope similarity (TPC percent similarity) than two having a 65% slope similarity.

While the TPC renders thermal pattern analysis more objective, it is still susceptible to subjectivity as it requires the examiner to use a best judgment approach when aligning the scans. If the TPC methods are found to be reliable, the next step would be to determine if the method has validity.

The Scan

The TyTron C-3000 [Titronics Research & Development, Oxford, IA] is a dual-probe infrared instrument that records temperatures on both sides of the spine and also performs a secondary calculation to determine bilateral temperature variations (delta). The TyTron instrument and protocol has been previously described elsewhere.[9, 14] A typical thermographic scan is begun at L5 and continues to the occipital shelf (Figure 1). The scanning procedure has been shown to have high reliability.[14] Each scan is comprised of three vertical lines (readings) displayed on the computer monitor and is comprised of one line for each side of the spine and a third line representing the delta (side-to-side differences) (Figure 2). These lines are also referred to as *channels*. Figure 3 shows a series of left paraspinal readings. In the past, the simple qualitative observation method of inspection was used to assess the amount of similarity (pattern) between the readings. This is inherently a subjective process, and any analytic techniques which minimize this subjectivity and maximize objectivity are clearly desirable.

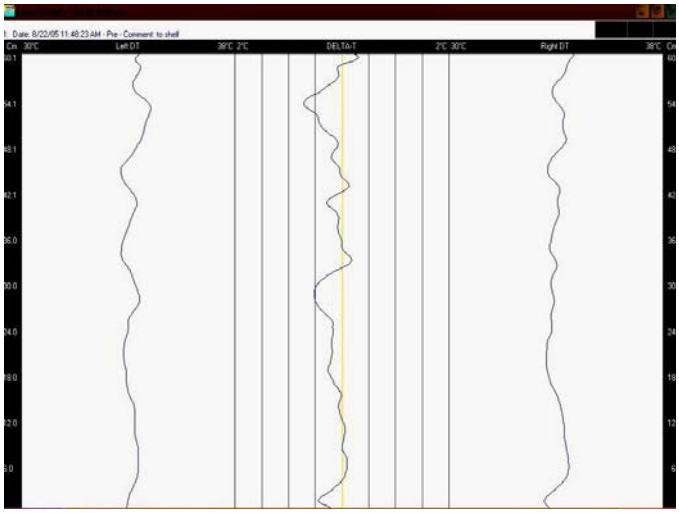
Ideally, a thermal scan stops at exactly the same point for every scan; however, in actuality this does not happen

Figure 1 *TyTron scanning procedure*.

due to examiner variations when attempting to stop at a given point on the spine for the purpose of comparing successive scans. Consequently, the readings being compared in the TPC do not line-up precisely (Figure 4). The TPC reads data points between any two scans from side-to-side, for 350 to 400 rows of temperatures on each side of the spine; therefore, it is important that the readings be vertically aligned as evenly as possible to achieve the most precise calculation (Figure 5). The process of vertical alignment can itself be a source of subjectivity. Another source of subjectivity of the analysis pertains to the extensions of the upper and lower data points in the graphic display (Figure 6). Extending the data points could include up to a maximum of 20 rows of temperatures, as part of 350 to 400 rows overall (Figure 5). Thus, the examiner has the option of leaving the data *in situ* or to slide the graphs to enhance vertical alignment as well as extending the scans from the top or bottom if the data are deemed to be useable. The sliding of the scans and extending of the data points are herein referred to as *modifications*.

Due to these options, the question arises as to whether the different approaches affect the outcome of the TPC analysis. Further to the findings of Owens et al., this study seeks to:

Figure 2 *TyTron scan results before being imported to the TPC software.**



Left Delta Right

**The line on the left represents the left paraspinal temperatures, the line on the right represents the right paraspinal temperatures and the middle line represents the delta or difference between the left and the right. The vertical lines within the delta channel represent scaled increments of one-half degrees.*

Figure 3 *Series of left paraspinal readings.*

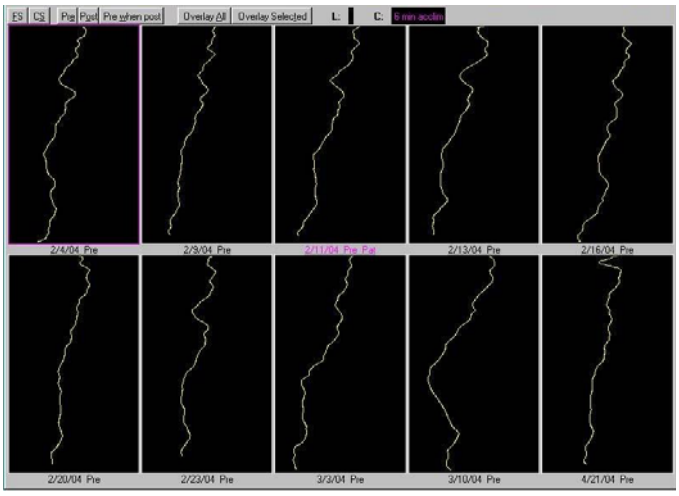
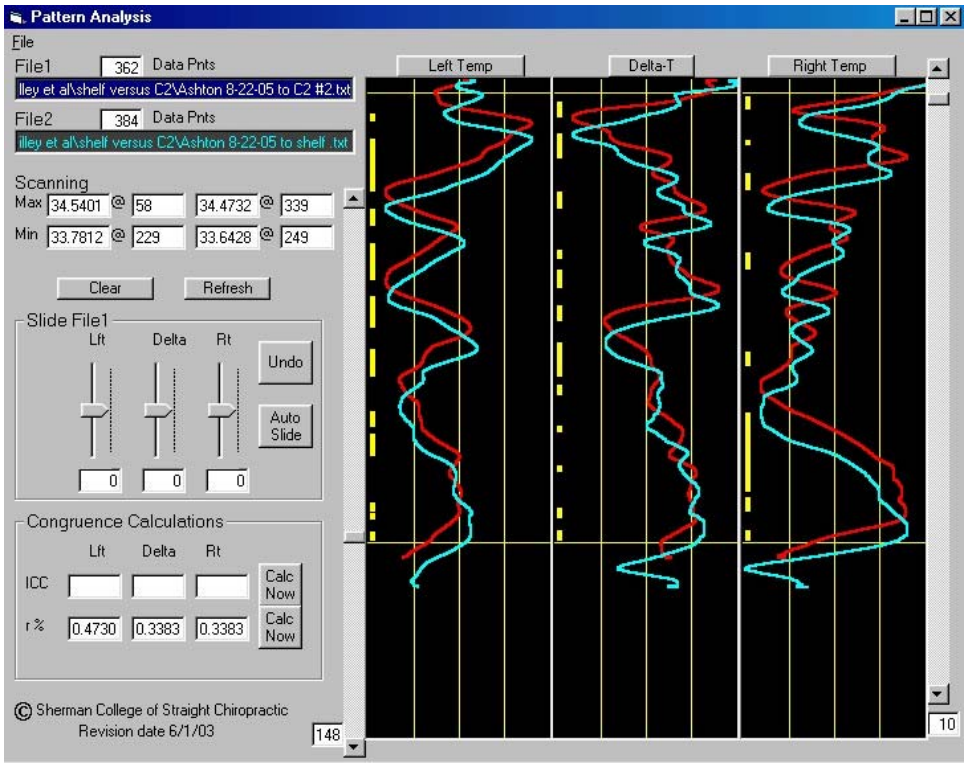


Figure 4 *Readings that do not line-up evenly due to operator error.**



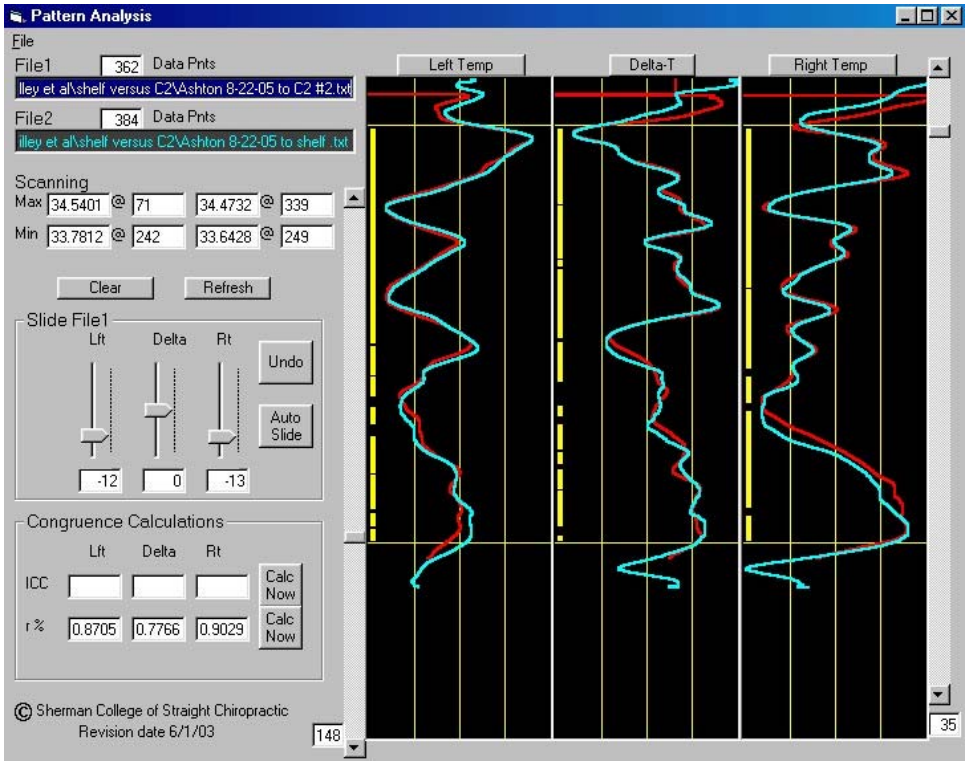
**This Figure also represents Method 3 TPC results. The same data points as depicted in Figures 5 and 6 are also presented in this Figure. Note that the data points depicted in graphic form do not match perfectly due to variations in stopping points when the thermographic scan was taken. It is apparent that usable data points above the upper bar and below the lower bar are present. However, this method did not allow for vertical alignment or modification of these data points as the graphs were left in situ. Also note the smaller percent agreement for the individual channels (r%, lower left on the TPC screen) compared to Methods 1 and 2 which allowed the examiner to vertically align the graphs to achieve a better match. The numbers 148 and 10 indicate the borders of the cropped area for the lower border (148) and upper border (10).*

(a) Assess the *intra-examiner reliability* for each of three thermal pattern calculation (TPC) methods;
(b) Assess the *inter-examiner reliability* for each of three TPC methods;
(c) Assess the by-examiner *inter-method differences* of the three TPC methods; and
(d) Determine which of the three TPC methods yields the *highest percent-similarity* between paired test-retest scans.

Methods

The study was approved by the Sherman College IRB and participants signed a consent form. Using a convenience sampling method,[15] 30 students (16 females, 14 males) were recruited from one of the authors' classes for thermographic scanning purposes. For feasibility reasons, we did not perform *random sampling;* nevertheless, we have no reason to believe that our sample should yield data so biased to the extent of diametrically changing our

Figure 5 *Method 2 TPC results.**



**The same data points as depicted in Figure3 are also presented in this Figure. The graphs have been vertically aligned, but no modifications of data points (above the upper bar or below the lower bar) have been included (arrows).*

conclusions. Participants' ages ranged between from 22–55 years (mean = 28.9 years, median = 26 yrs).

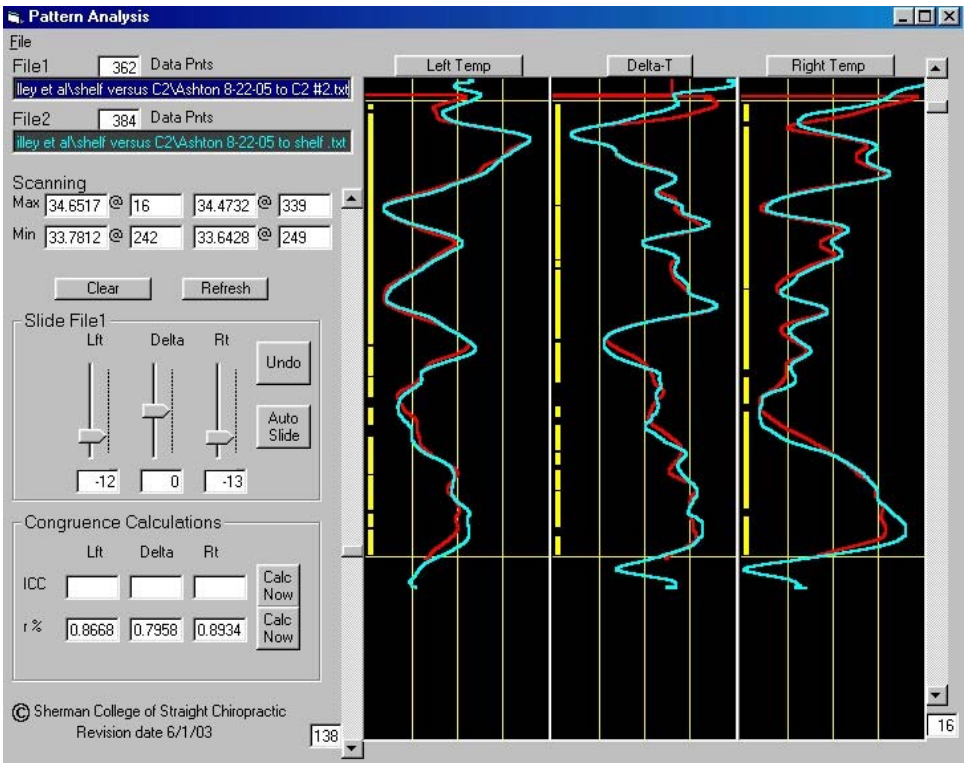
Each volunteer was scanned twice with the TyTron C–3000, with 10 minutes between the scans. The scans were performed by the principal investigator who had approximately six years of thermal scanning experience with this thermal instrument. From the TyTron program, the readings were then exported into a notepad (text) file consisting of several hundred rows of numerical temperatures for each reading. The longer the length of the participant's spine, the more rows of temperatures there were. The notepad files were then imported into the TPC software where the two readings were graphically displayed and compared for similarity of slope as previously described.[12]

A convenience sample of three independent examiners was selected from available research department employees. One of the examiners had approximately four years experience working with the TPC program, while the other two examiners each had approximately one month of experience using the TPC. A brief training session was given by the principal investigator for the other two examiners for the purpose of establishing consistency of procedures. Each examiner compared 60 thermal scans from 30 different participants (two scans per participant).

TPC Methods 1–3

Method 1 (Figure 6) involved vertically aligning the two scans from each of the 30 participants. When a scan is vertically aligned, the entire scan is moved for a best fit according to the examiner's visual assessment. Once the examiner is satisfied, the "calculate" button is clicked and the TPC percent similarity (of slope) is returned. A second procedure of deleting or including extra data points

Figure 6 *Method 1 TPC results.**



**Data points have been vertically aligned by the examiner to visually achieve the best match for the left and right side graphs. Modification of extra data points above the upper bar and below the lower bar, to include usable data, has been incorporated as well. If the examiner chooses to include data points above the upper bar or below the lower bar, as was done with Method 1, the amount of extra data points always defaults to the shorter graph. The darker reading is longer than the lighter reading (arrow) due to different lengths of scans (arrows).*

(Figure 5), if deemed usable, was also a step in the protocol. Method 2 involved using the same amount of vertical alignment as in Method 1 but did not include the option of modifying extra data points (Figure 5). Method 3 (Figure 4) was entirely computer generated; the only requirement of the examiner was to press the "*calculate*" key for the TPC software to calculate the percent similarity.

The examiners performed Method 1 in the first session, Method 2 in a second session and Method 3 in a third session. The sessions ranged between from one to three weeks apart. It was considered highly unlikely that the examiners would recall how they had previously aligned or modified data on 30 sets of graphs from the previous session. However, if an examiner did recall any aspects of performing a prior method, that recall could possibly have influenced performance of the next method. Consequently, the precaution was taken to have each examiner attest, in writing, to their level of recall regarding the degree of aligning or modifying data that had transpired during the previous session. All reported they had no recall of how they had manipulated the scanned data while performing Method 1 or Method 2. As an additional measure to protect against any recognition of data, each thermograph scan was given a different identification number for each of the two trials and the sequence of the numbers changed as well. Recall was not

an issue in Method 3, as that method was entirely computer driven, without any input from the examiners in regard to aligning or otherwise modifying data. Thus, each examiner was blinded to results of his/her previous session(s) as well as to any of the results of the other examiners. The examiners made hand-written percent notations for their respective calculated percent agreements of slope between the scans being compared in the TPC software. The examiners were assessed as to whether they each agreed on the resultant percent similarity returned by the TPC software after necessary alignment and optional data points procedures were addressed. These results were then manually entered into a spreadsheet for statistical analyses.

Statistics

Since the examiners were not randomly selected, the Fixed Effects (Model 3) Intraclass Correlation Coefficient (ICC) was used[16] to determine the extent of intra- and inter-examiner agreement for all three methods. The minimum acceptable ICC score was set at 0.750[16] with an alpha level of 0.05 (5% level of significance). The ICC test assessed the TPC percents (slope similarity) obtained by the examiners.

The Kolmogorov-Smirnov and Shapiro-Wilk statistics demonstrated that the data were not normally distributed. Consequently, the non-parametric Wilcoxon signed-rank test, which does not assume normal distribution, was used to assess intra- and inter-examiner differences as well as differences between methods. The alpha level for the Wilcoxon test was set at 0.05 (5% level of significance) The ICC, Kolmogorov-Smirnov, Shapiro-Wilk, and Wilcoxon signed-rank tests were performed with the Statistical Package for the Social Sciences [SPSS Version 14.0, Chicago, IL].

Results

TPC methods

Visual depictions of the three methods are shown in Figures 4–6. The same thermographic temperatures (data points) are represented in these three figures, with explanations provided for each of the methods.

Examiner agreements

Method 1 intra-examiner reliability (repeatability) ICC scores ranged from 0.860 to 0.987 (Table 1). Inter-examiner ICC scores for this method ranged from 0.799 to 0.926 (Table 2). Method 2 intra-examiner reliability ICC scores ranged from 0.791 to 0.982 (Table 1). Inter-examiner reliability ICC scores for this method ranged from 0.805 to 0.941 (Table 2). Because Method 3 required no modifications to the graphs, intra and inter-examiner comparisons naturally did not reveal any significant differences and ICC scores were a perfect 1.00, reflecting the reproducibility of the TPC software (Tables 1 & 2).

Examiner differences

There were no significant differences for intra and inter-examiner agreement for any of the methods (Tables 1 and 2). There were 14 significant differences found (out of 36 comparisons) when comparing Method 3 to Methods 1 and 2 (Table 5). Methods 1 and 2 produced higher percentages of similarity when comparing paired test-retest graphs than Method 3 (Table 5). As shown in Table 5, there was a small, clinically trivial difference in percent-similarity between Methods 1 and 2, varying from 0.3 to 0.4, while a substantially larger difference in percent-similarity occurred between Methods 1 and 3 (4.0 –7.1) and Methods 2 and 3 (4.3 to 7.5).

Discussion

Initially, in the present study, there was interest to investigate if a difference in experience using the TPC software would affect the outcome. Only one of the three examiners had considerable experience with the TPC software. The second examiner was a chiropractic student while the third was a chiropractic technician, both of whom had approximately one month of experience with the TPC. Thus, variation of technical training with the TPC software was considerable; however the high degree of repeatability of the three examiners suggests that the TPC procedures are readily learned, thereby potentially providing a valuable technology for a variety of practice settings.

The results show that different examiners can demonstrate a high level of within-method agreement, regardless of which of the three tested methods they apply; nevertheless, although Methods 1 and 2 maximized percent-similarity by affording the examiner the option of visually aligning the scanned data before importing to the TPC, and indicated good inter-method agreement with

Table 1 *Three Methods*[1] *Intra-examiner ICC*[2] *Scores and P-values*[3]

	Examiners								
	A			B			C		
	Channels								
	Left	Delta	Right	Left	Delta	Right	Left	Delta	Right
Method 1									
ICC	0.987	0.959	0.976	0.892	0.860	0.892	0.952	0905	0.968
p-value	0.7	0.8	0.9	0.7	0.9	0.9	0.7	0.6	0.8
Method 2									
ICC	0.980	0.969	0.982	0.880	0.791	0.913	0.910	0.824	0.975
p-value	0.9	0.9	0.9	0.4	0.8	0.9	0.9	0.7	0.8
Method 3									
ICC	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
p-value	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

1. Thirty subjects each had two thermograhic scans taken 10 minutes apart. Three examiners twice compared the 30 pairs of scans by three different methods. Each method was applied in separate sessions (see Methods).
2. Intraclass Correlation Coefficients. All ICC scores exhibited the minimum acceptable level of 0.750 or greater. Method 3 was perfectly correlated as it involved no input from the examiner other than to press "calculate" on the computer's keyboard. There were no significant differences between trials 1 and 2 for any of the three methods.
3. P values represent difference according to the Wilcoxon test (alpha= 0.05) and was used to assess differences between Trials 1 and 2.

each other, one-third of their results differed significantly from Method 3 ($p < 0.05$), the latter of which required no input from the examiner (Table 4).

Limitations of the Study

Because of the many multiple comparisons in the study, the original 5% level of significance (alpha) was subjected to a Bonferroni correction that considerably lowered the alpha (to 0.0004). This in turn, resulted in none of the inter-method comparisons being able to demonstrate any statistically significant differences (where with the traditional alpha of 0.05, there were 14), which suggests that some of our analyses may be subject to a Type 2 error (accepting an erroneous hypothesis).[17] There is controversy as to whether Bonferroni adjustment should be used[18] and consensus is lacking as to when it is appropriate.[19]

The study participants do not represent a random sample from the general population (they were generally healthy, young adult students), so that the generalizability of our findings may be diminished; however, because there was a wide variation of thermal patterns (an important criterion in reliability testing), we judge our sample profile to be adequate for the purposes of our reliability analyses.

An ICC for inter-method analysis was not performed because of the differing protocols required in the different methods; thus it was believed that there would not be an apple-to-apples comparison being made which was thought to be a violation of ICC calculations.

These findings are preliminary and should be verified with further research. Furthermore, in light of the growing body of evidence demonstrating the reliability of the scanning method as well as to the TPC software, other validation aspects of thermal pattern analysis now require evaluation. While this study contributes information in regard to examiner reliability in applying more than one method of importing thermal scan data into the TPC software, as well as how those methods affect the TPC re-

Table 2 *Three Methods*[1] *Inter-examiner Agreement (ICC)*[2]

	Intra-class Correlation Coefficients (ICCs)		
	Left	Delta	Right
Method 1			
Trial 1	0.829	0.799	0.918
Trial 2	0.910	0.881	0.926
Method 2			
Trial 1	0.815	0.806	0.941
Trial 2	0.825	0.805	0.936
Method 3			
Trial 1	1.00	1.00	1.00
Trial 2	1.00	1.00	1.00

1. Thirty subjects each had two thermographic scans taken 10 minutes apart. Three examiners twice compared the 30 scans by three different methods. Each method was applied in separate trials (see Methods). All ICC scores exhibited the minimum acceptable level of 0.750 or greater. Method 3 was perfectly correlated as it involved no input from the examiner other than to press "calculate" on the computer's keyboard.
2. Intraclass Correlation Coefficients. All ICC scores exhibited the minimum acceptable level of 0.750 or greater. Method 3 was perfectly correlated as it involved no input from the examiner other than to press "calculate" on the computer's keyboard. There were no significant differences between trials 1 and 2 for any of the three methods.

Table 3 *Inter-examiner Differences for Three Methods*[1]

Examiners	Left Channel Trial 1 / Trial 2	Delta Channel Trial 1 / Trial 2	Right Channel Trial 1 / Trial 2
Method 1			
A-B	0.9 / 0.5	0.8 / 0.6	0.6 / 0.8
A-C	0.4 / 0.4	0.2 / 0.4	0.4 / 0.6
B-C	0.4 / 0.9	0.3 / 0.6	0.6 / 0.8
Method 2			
A-B	0.9 / 0.3	0.9 / 0.6	0.8 / 0.8
A-C	0.2 / 0.2	0.2 / 0.3	0.4 0.6
B-C	0.2 / 0.8	0.3 / 0.6	0.6 / 0.7
Method 3			
A-B	1.0 / 1.0	1.0 / 1.0	1.0 / 1.0
A-C	1.0 / 1.0	1.0 / 1.0	1.0 / 1.0
B-C	1.0 / 1.0	1.0 / 1.0	1.0 / 1.0

1. The numbers provided represent the p value of the Wilcoxon signed-rank test (alpha = 0.05 [Bonferroni corrected = .0004]) for inter-examiner differences across each trial separately. Thirty subjects each had two thermograhic scans taken 10 minutes apart. Three examiners twice compared the 30 scans by three different methods. Each method was applied in separate sessions (see Methods).

Table 4 *Inter-method Differences by Examiner and Trial*[1]

		Method								
		1 vs 2			1 vs 3			2 vs 3		
Examiner	Trial	Left	Delta	Right	Left	Delta	Right	Left	Delta	Right
A	1	0.7	0.7	0.8	0.1	0.02	0.04	0.05	0.01	0.05
	2	0.8	0.6	0.9	0.07	0.01	0.04	0.04	0.006	0.04
B	1	0.6	0.6	0.8	0.1	0.03	0.06	0.05	0.01	0.06
	2	0.9	0.8	0.8	0.2	0.02	0.1	0.2	0.02	0.07
C	1	0.8	0.7	0.8	0.3	0.1	0.2	0.3	0.1	0.2
	2	0.9	0.9	0.9	0.2	0.04	0.1	0.2	0.04	0.1

1. Each examiner twice compared two graphs from 30 subjects for analysis of percent similarity. There were no significant differences between the methods.

Table 5 *TPC Means and Standard Deviations of Percent Similarity for Thirty Pairs of Thermographs*[1]

	Left	Delta	Right
	Percent Similarity		
Method 1			
Mean	72.4	70.0	72.3
Std. Dev.	7.8	8.1	9.5
Method 2			
Mean	72.7	70.4	72.7
Std. Dev.	7.7	8.6	9.8
Method 3			
Mean	68.4	62.9	67.3
Std. Dev.	10.1	12.7	11.1
	Differences in Percent Similarity Between the Three Methods		
Methods 1:2	0.3	0.4	0.3
Methods 1:3	4.0	7.1	5.0
Methods 2:3	4.3	7.5	5.4

1. Thirty subjects each had two thermograhic scans taken 10 minutes apart. Three examiners twice compared the 30 pairs of scans by three different methods. Each method was applied in separate trials (see Methods).

sults, we recommend that the next step towards validating this thermal pattern analysis technique would be to assess its validity.

Conclusions

The TPC software program itself was shown in our study to be acceptably reliable. Aligning the readings, as used in Methods 1 and 2, appears to play an important role in maximizing TPC percent-similarity, and both of these methods descriptively demonstrated acceptable within- and between-method reliability. (Recall that Method 2, incorporated the same amount of vertical aligning as Method 1, but omitted the extra data inclusions.)

Our study also provides evidence that the key TPC procedures we used are easily learned and reliable. The findings of this study prompted the initiation of improvements to the TPC program which enables the software to automatically align thermographs, thereby eliminating another area of examiner subjectivity in the use of the TPC. Future studies should involve: a) a random selection process for participants (from the general outpatient population) and examiners in order to better generalize the findings, (b) comparing Methods 1 and 2 with software that automatically aligns graphs for analysis, and c) investigating the relationship between TPC findings and other indicators of vertebral subluxation (construct validity).

Conflict of Interests

None declared

References

1 Chiropractic Paradigm. Association of Chiropractic Colleges. Retrieved November 4, 2005 at http://www.chirocolleges.org/paradigm_scopet.html

2 Uematsu S, Edwin DH, Jankel WR, Kozikowski J, Trattner M. Quantification of thermal asymmetry—part 1; normal values and reproducibility. J Neurosurgery 1988; 69:552–555.

3 Kim YC, Bahk JH, Lee SC, Lee YW. Infrared thermographic imaging in the assessment of successful block on lumbar sympathetic ganglion. Yonsei Med J. 2003; 44(1):119–24.

4 Owens EF, Pennacchio VS. Operational definitions of vertebral subluxation: A Case Study. Top Clin Chiropr 2001; 8(1):40–48.

5 Guyton AC, Hall JE. Textbook of Medical Physiology. 9th Edition. Philadelphia: Saunders. 1996. p. 912.

6 Palmer BJ. Chiropractic Clinical Controlled Research. Hammond, IN: W.B. Conkey; 1951 pp. 361–476.

7 Hart J. Skin temperature patterns of the posterior neck used in chiropractic analysis. Chiropractic: Journal of Chiropractic Research and Clinical Investigation 1991; 7(2):46–48.

8 Kessinger RC, Boneva DV. Vertigo, tinnitus, and hearing loss in the geriatric patient. J Manipulative and Physiol Ther 2000; 23(5):359–360.

9 Hart JF, Owens EF. Stability of paraspinal thermal patterns during acclimation. J Manipulative and Physiol Ther 2004; 27(2):109–117.

10 Stewart MS, Riffle DW, Boone WR. Computer-aided pattern analysis of temperature differentials. J Manipulative Physiol Ther 1989; 12(5):345–352.

11 Hart JF, Boone WR. Pattern analysis of paraspinal temperatures: a descriptive report. J Vert Sublux Res 1999–2000; 3(4):1–8.

12 Owens EF, Stein T. Computer-aided analysis of paraspinal thermographic patterns: a technical report. Chiropr Res J 2000; 7(2):65–69).

13 Owens EF, Hart JF, Stein T. Short-term stability and reliability of paraspinal infrared thermal scans. Ninth Annual Vertebral Subluxation Research Conference. October 13–14, 2001. Spartanburg, SC.
14 Owens EF, Hart J, Donofrio J, Haralambous J, Mierzejewski E. Paraspinal skin temperature patterns: An interexaminer and intraexaminer reliability study. J Manipulative Physiol Ther 2004; 27:155–159.
15 Baumgartner TA, Strong CH, Hensley LD: Conducting and Reading Research in Health and Human Performance. New York, NY: McGraw-Hill; 2002, p. 133.
16 Portney LG, Watkins MP. Foundation of Clinical Research; Applications to Practice. Upper Saddle River, NJ: Prentice Hall Inc; 2000, pp. 562–565.
17 Definition of type II error. Everything Bio. Retrieved January 19, 2007 at: http://www.everythingbio.com/glos/definition.php?word=type+II+error
18 Perneger TV. What's wrong with Bonferroni adjustments. BMJ 1998, 316: 1236–1238. Retrieved August 9, 2006 at: http://bmj.bmjjournals.com/cgi/content/full/316/7139/1236?view=full&pmid=9553006
19 Nakagawa S. A farewell to Bonferroni: the problem of low statistical power and publication bias. Behav Ecol 2004;15 (6):1044–1045.