MULTIVARIATE ANALYSIS OF HAND USAGE ON STRUCTURE AND FUNCTION

Dennis B. Gillings, Diane Makuc, Nortin M. Hadler
University of North Carolina at Chapel Hill

Summary

Three groups of female textile workers, each employed in different well-defined repetitive manual tasks for at least 20 years were identified in a single rural mill. Replicate data were obtained at one point in time for the following measurements on both hands: range of motion, degree of degenerative joint disease at each joint, malalignment at digital joints, osteophyte formation. Data items were either continuous or ordered categorical in nature and the joints of both hands for the same individual provided a multivariate profile of measurements.

Multivariable categorical data methods, and multivariate non-parametric and parametric techniques were employed to determine 1) observer agreement; 2) right and left hand differences; and 3) task differences. The statistical tests that were used are described.

It was concluded that there was 1) adequate observer agreement; 2) significant and consistent differences between the dominant right hand and the left; and 3) significant task related differences that were consistent with the pattern of usage in the industrial setting.

DESIGN OF STUDY

Selection and Description of Subject Groups. A single long established worsted mill in a small rural Virginia town was chosen. The mill employs over 600 people and is a major employer in a rural community with a stable population. It is characteristic of such plants that inter-job mobility is not commonplace. Three different manual tasks, burling, spinning, and winding that employ a large percentage of the workers were chosen because the task description has changed little over the past decades. Spinners and winders tend two different types of machines which (1) spin crude yarn into tighter thread and (2) wind several of the spun threads together for weaving. Burlers repair defects in the woven cloth.

Table 1 lists some characteristics of the study groups by task. Only female employees working continuously in the respective tasks for at least 20 years were considered eligible. Of those eligible, the number of volunteers is listed. No attempt was made to determine the motivations of those who elected not to participate.

Execution of the Clinical Study. The study was executed over the course of three working days. Subjects in groups of four were scheduled to arrive at 30 minute intervals at the on-site conference rooms made available by plant management. In a private room, they were individually interviewed by the research secretary and assigned a study number. The interview entailed completing a questionnaire and then reading and signing a consent form. During the remainder of the clinical exam, subjects were identified only by study number.

TABLE 1. DESCRIPTION OF STUDY GROUPS BY TASK

	Burl	Wind	Spin
Eligible Employees	39	16	20
Volunteers	32	16	19
Unavailable*	2	0	0
Excluded†	1	0	0
Total in Study	29	16	19

*Hospitalized with intercurrent illness.

†Inflammatory polyarthritis was detected in the course of examining this subject.

Clinical Examination. There were four examiners: three rheumatologists and a physical therapist. This team was repeatedly rehearsed prior to the actual study. The four subjects were randomly assigned to each of the four examiners. The examination required 15 minutes after which the same subjects were randomly assigned to a second examiner different from the first. Therefore, each subject in the study was examined by two separate investigators. The random assignment schedule was designed prior to the study.

Three categories of data were obtained by clinical examination:

1. The hands were examined for the presence of synovitis or the residua of major overt trauma. One subject (Table 1) was excluded from the study because of inflammatory synovitis and is currently under continuing management. Five subjects were detected who had incurred major trauma in the distant past with residual deformity to one digit.

2. The extremes of active range of motion were measured to the nearest 15 degrees with a small plastic universal goniometer. Fifteen degrees was chosen as the categorical unit of measurement as differences of this magnitude are less subject to doubts as to clinical significance. All small joints of the hand, the first carpometacarpal joint and the wrist were examined. An example of the data form used throughout the study is shown in Figure 1.

3. The circumferences of all distal (DIP) and proximal (PIP) interphalangeal joints were determined with an arthrocircameter.

Radiographic Examination. Radiographs were taken by the office technologist in the employ of an orthopedic surgeon in the community. Subjects were transported by the plant nurse in small groups in the weeks prior to the clinical study. A single postero-anterior radiograph of both hands and wrists was obtained. The radiographs were identified only by a numerical coded marker. Each radiograph was evaluated by two rheumatologists without access to the code. Each joint for which goniometric data was obtained was scored for DJD from 0 to 4. In addition, the minimal width of the mid-phalangeal shaft was recorded to the nearest millimeter. Malalignment was determined by measuring, to the nearest 5°, the discrepancy in the alignment of the long axes of the contiguous bony shafts at the PIP and DIP joints, excluding the thumb.

Description of the Tasks. The task description was performed by Dr. M.A. Ayoub, a consulting industrial engineer and ergonomicist. A micromotion analysis was not performed. A standard time-motion analysis was made available by the plant industrial engineer. Dr. Ayoub supplemented this information with direct observation. Without access to the results of the study, he responded to a set of direct questions formulated by the investigators in order to rank the tasks by frequency of patterns of hand use.

STATISTICAL METHODS AND RESULTS

Description of the Groups of Workers and Their Tasks. The questionnaire was designed to provide some demographic data and to assess the homogeneity of the groups by task. The results are presented in Table 2. There were no significant task differences. Almost all subjects had lived their entire lives within the contiguous counties. Only two pairs of subjects were relatives. The groups did not differ in age (Table 2: Kruskal-Wallis test $P = 0.1769$). Further, there was no evidence of self-selection on the part of individuals as regards choosing which job to perform. Few persons admitted that they requested a position of spinning, burling, or winding (see Table 2).

It is apparent from Table 2 that the subjects perceived only winding as a bimanual task. All tasks are highly repetitive, stereotyped, and complex in that to some extent they are bimanual. However, in terms of frequency, the task analysis corroborates the perceptions of the subjects. Furthermore, winding differs distinctly from the other two because of a predominance of wrist motion and the employment of a power grip with little fine finger motion. Burling and spinning differ in that the latter tends to utilize a three-finger hand, sparing in use digits 4 and 5. These task descriptions were patently obvious results of the assessment undertaken. A more detailed description is not justified in the absence of an extensive micromotion analysis quantifying the force employed and frequency of use at each joint.

Data Analysis. Four separate response variables resulted from the clinical and radiographic studies: range of motion, malalignment, radiographic DJD score, and derivatized circumference. All data were obtained independently by two observers. The data analysis was planned to answer three main questions:

a) Were the observers in agreement as regards the data items recorded on each individual?
b) Were there differences between the right and left hands of each individual?
c) Were there differences between individuals who worked on different tasks?

The strategies that were determined to answer each of these questions are outlined below. Joint groups with trauma-induced deformity (five) and the two miscellaneous isolated missing values were excluded from the corresponding multivariate analyses. That two subjects considered their task as left-handed was disregarded except where indicated.

Analysis of Observer Agreement. The analysis of observer agreement was carried out to test agreement between observers as regards the measurement of individual subjects.

A descriptive assessment of observer agreement was obtained by computing the percentage of individuals with perfect agreement for both observers and the percentage of individuals with agreement for both observers within one unit of measurement. The degree of perfect agreement is presented in Table 3.

The measuring instruments were not really designed to have high perfect agreement, especially with respect to range of motion where measurements were taken within 15°. The percentages of individuals who show agreement within one unit of measurement are invariably high (data not shown), the majority being over 90%. Only 3 such percentages fell below 80%. The best agreement within one unit of measurement (100%) for range of motion is found in joints 9 through 12, and the worst, (71.4%) in joint 13 for the left hand.

Kappa-type statistics as described by Landis (1) provide a more refined measure of the extent to which two observers classify individual subjects into the same category. Weights are used to vary the definition of agreement. Kappa statistics were computed for each of the four response variables on joints with crude agreement less than 90%. Beyond 90%, agreement was felt to be more than adequate. For this analysis two sets of weights were considered, one corresponding to perfect agreement ($\hat{k}_1$) and the second to agreement within one measurement unit ($\hat{k}_2$). Kappa statistics can range from 0.00 to 1.00 with values from 0.00 to 0.20 roughly indicating "slight" agreement and 0.81 to 1.00 "almost perfect" agreement.

The data for right and left hands were considered independent in this analysis. All results were obtained with the computer program GENCAT (2) which is a general routine that analyzes categorical data by weighted least squares according to the formulation developed by Grizzle, Starmer, and Koch (3). The kappa statistics are presented in Table 4. The lowest value of $\hat{k}_1$ was 0.093 for range of motion at joint 7 and the highest was 0.538 for derivatized circumference at joint 13. All except three of the kappa statistics were significantly different from zero at the .05 level. For those measurements with agreement within one unit of measurement less than 90%, values of $\hat{k}_2$ ranged from 0.272 to 0.712. All of the kappa statistics for derivatized circumference and malalignment were significantly different from zero at the .01 level. The analysis using kappa statistics supports the conclusion derived from the measures of crude reliability, namely that agreement between the first and second observers is acceptable.

Differences Between Right and Left Hands. Multivariate sign tests were applied where the right hand was ranked 1 if it had the greater score or measurement and 2 if it had the smaller. A mid-rank 1.5 was assigned if the right and left hands were equal. This test is based on the multivariate version of Friedman's test (4) as described by Gerig (5). The test compares mean ranks for each hand simultaneously across a profile of joints. Computations were carried out using the computer program FLOTA (2). The multivariate version of Friedman's test is equivalent to a multivariate sign test in the case where there are only two treatments (treatment corresponds here to right or left hand) in the same way that Friedman's test is equivalent to a (univariate) sign test when there are only two treatments.

The multivariate sign test was applied separately to each of the five digits and to the wrist. The results of the tests of significance are summarized schematically in Figure 2. It can be seen that there are extensive right and left hand differences. The multivariate test was significant in each case where data was available for range of motion and derivatized circumference. Several of the corresponding univariate tests were also significant, thus supporting these findings. None of the multivariate tests were significant for DJD score, although two of the univariate tests reached significance at the 0.05 level. There were no significant right and left hand differences in respect to malalignment.

Difference Between Tasks. Task differences were assessed in four different ways, referred to as Models 1-4.

Model 1 - right task hands only considered as the unit of observation.

Model 2 - left hands only considered as the unit of observation.

Model 3 - right and left hands considered separately, each hand being taken as an independent unit of observation.

Model 4 - right and left hands jointly considered as the unit of observation.

For each of Models 1-4 a multivariate Kruskal-Wallis test, as discussed by Koch (6) was applied to each profile of scores for the range of motion, malalignment and radiographic DJD scores. The means of the measurements for the two observers were the data points analyzed. The null hypothesis was that there were no differences among tasks. If the null hypothesis is rejected, then at least one of the tasks is different from the others. The direction of such differences may be clarified by inspection of the corresponding mean ranks. Joints were considered separately for univariate analysis and grouped by digit and wrist for multivariate analysis.

For this analysis, the treatments were taken to be the three tasks spinning, burling, and winding. The scores for each joint were combined across tasks and ranked. The multivariate Kruskal-Wallis test was used to compare tasks simultaneously across joints for each digit and the wrist using the computer program SPLOTA (2). Corresponding univariate Kruskal-Wallis tests (7) for each joint were routinely computed by SPLOTA to facilitate greater understanding of any differences detected. The results for this analysis and the MANOVA analysis described below are summarized in Figure 3.

The multivariate Kruskal-Wallis test was also appropriate for the derivatized circumference of joints. However, a MANOVA (8) analysis was selected as the procedure of choice as the data were continuous. Age was included as a covariate. The analysis was performed with task as the main effect. The null hypothesis was again tested for each of the models, right task hands, left hands, right and left hands separately, and both hands as the unit of observation. Hands with injured joints or missing data were eliminated from the analysis.

The most striking findings are summarized in Figure 3. Only p-values that were consistent in the several modes of testing including the task analysis, are presented in this figure. There is a problem of multiple comparisons with the large numbers of statistical tests and so consistency of results was chosen as the approach to handle this.

There were several task-related differences that were statistically significant for the left hand but not included in Figure 3 since these left-hand task-related differences were clinically unimportant resulting from differences within a single unit of measurement.

DISCUSSION

This study is cross-sectional and retrospective in design. It therefore suffers from the inherent flaw that we are unable to comment on the loss from the cohort that was first employed 20 years ago. Nonetheless we detect multiple and consistent differences in the structure and function of the hands of the women employed in the three tasks. It is our contention that these differences argue cogently for the rejection of the null hypothesis that pattern of usage does not influence structure and function of the hands. Attrition from the initial cohort is certainly multifactorial. However, attrition because of hand DJD would be expected to obscure the task related differences we observed.

Further, right-left differences were readily detected in the analysis of range of motion although such differences were demonstrable when derivatized circumference and DJD scores were analyzed (Figure 2). These demonstrations of greater impairment in the right hand by themselves argue for a role of usage (a traumatic element (9)) in the pathogenesis of primary DJD of the hands.

It is important to realize that this study is not designed to test for abnormality. We are testing the likelihood that a single independent variable--the pattern of usage--influences the structure and function of the hands of three groups of women that are highly comparable. To test for abnormality one would need to identify a control group that was normal in terms of structure and function. The ideal control group would preferably lack the influence of the independent variable under study. Since that variable is hand usage, there is no ideal control. An alternative control group would be one in which all possible patterns of hand usage are represented without bias. If these patterns were defined and their frequency of occurrence in a suitable population known, then the usage patterns in a control group could be measured to demonstrate the absence of bias. However, it would be prohibitively expensive to identify a control group in this way. An alternative would be to take a simple random sample of middle aged women from a large population base. This would still have been potentially uninformative as no difference between experimental and control groups would not rule out the possibility that usage affects structure and function. Several types of usage would occur in the control group and some of these could affect structure and function more than others. One is then led to conduct a study in which the independent variable is manipulated in a more precise way. The above is an example of such a study.

REFERENCES AND NOTES

1. Landis, J.R. and G.G. Koch. 1977. The measurement of observer agreement for categorical data. Biometrics 33:159-174.

2. Development of GENCAT, FLOTA, SPLOTA computer programs, coordinated by Dr. Gary Koch, Dept. of Biostatistics, School of Public Health, University of North Carolina, Chapel Hill, North Carolina 27514.

3. Grizzle, J.E., C.F. Starmer, and G.G. Koch. 1969. Analysis of categorical data by linear models. Biometrics 25:489-504.

4. Friedman, M. 1973. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J. Am. Stat. Assoc. 33:675-699.

5. Gerig, T. 1969. A multivariate extension of Friedman's test. J. Am. Stat. Assoc. 64:1595-1608.

6. Koch, G.G. 1970. The use of non-parametric methods in the statistical analysis of a complex splitplot experiment. Biometrics 26: 105-128.

7. Kruskal, W.H. and W.A. Wallis. 1952. The use of ranks in one criterion variance analysis. J. Am. Stat. Assoc. 47:583-621.

8. Smith, H., R. Gnanadesikan, J.B. Hughes. 1962. Multivariate analysis of variance (MANOVA). Biometrics 18:22-41.

9. Acheson, R.M., Y.K. Chan, and A.R. Clemett. 1970. New Haven survey of joint diseases. XII: Distribution and symptoms of osteoarthrosis in the hands with reference to handedness. Ann. Rheum. Dis. 29:275-286.

TABLE 2

SPECIFIED CHARACTERISTICS BY TASK

Specified Characteristic	Burl	Wind	Spin	Total
Hand dominance				
right	28(96.6%)	15(93.8%)	17(89.5%)	60(93.8%)
left	1(3.4%)	1(6.3%)	2(10.5%)	4(6.3%)
Hand performing the task				
right	28(96.6%)	1(6.3%)	19(100.0%)	48(75.0%)
left	1(3.4%)	1(6.3%)	---	2(3.1%)
both	---	14(87.5%)	---	14(21.9%)
Task always done in the same way	29(100.0%)	6(37.5%)	18(94.7%)	53(82.8%)
Never lived outside county	18(62.1%)	11(68.8%)	14(73.7%)	43(67.2%)
A relative also worked at least 20 years burling, spinning, winding	1(3.4%)	3(18.8%)	2(10.5%)	6(9.4%)
Serious injury to hand				
right	3(10.3%)	1(6.3%)	3(15.8%)	7(10.9%)
left	2(6.9%)	3(18.8%)	1(5.3%)	6(9.4%)
At least five years at one other job	2(6.9%)	1(6.3%)	---	3(4.7%)
Have hobby with repeated manual work	5(17.2%)	1(6.3%)	---	6(9.4%)
Spend more than five hours/week at manual hobby	2(6.9%)	1(6.3%)	---	3(4.7%)
Spent more than five years at manual hobby	3(10.3%)	---	---	3(4.7%)
Currently have manual hobby	3(10.3%)	1(6.3%)	---	4(6.3%)
Requested position in burling, winding, or spinning	4(13.8%)	1(6.3%)	1(5.3%)	6(9.4%)
Number of cases	29(100.0%)	16(100.0%)	19(100.0%)	64(100.0%)
Age (mean ± S.D.)	56.2 ± 7.7	49.0 ± 6.1	49.4 ± 6.0	

TABLE 3

PERCENT OF SUBJECTS WITH PERFECT AGREEMENT BETWEEN OBSERVATIONS

	Range of Motion		Derivatized Circumference		Malalignment		DJD Score	
Joint*	Right	Left	Right	Left	Right	Left	Right	Left
1	56.3	60.9					96.9	92.2
2	38.7	47.6					79.7	89.1
3	37.5	34.4					75.0	73.4
4	38.1	53.1					92.1	92.2
5	46.9	48.4					73.0	78.1
6	54.7	53.1					65.6	79.7
7	50.0	40.6					87.5	89.1
8	42.2	50.0					62.5	75.0
9	82.5	87.3	71.4	60.3	76.2	76.2	68.3	50.8
10	93.8	93.8	71.9	81.3	68.8	79.7	67.2	67.2
11	93.8	87.5	76.6	89.1	59.4	65.6	68.8	71.9
12	88.9	90.6	61.9	75.0	71.4	64.1	69.8	65.6
13	34.4	36.5	73.4	84.1			50.0	61.9
14	73.0	75.8	74.6	66.1	60.3	54.1	49.2	59.0
15	76.6	74.6	75.0	84.4	84.4	57.1	54.7	74.6
16	79.7	82.8	76.6	87.5	64.1	45.3	46.9	56.3
17	76.2	79.7	63.5	73.4	68.3	59.4	49.2	67.2
Number of hands	64	64	64	64	64	64	64	64

*Numbering of joints is explained in Figure 1.

TABLE 4

KAPPA STATISTICS* FOR RANGE OF MOTION, DERIVATIZED CIRCUMFERENCE, MALALIGNMENT, AND DEGENERATIVE JOINT DISEASE SCORE BY JOINT

Joint	Range of Motion $\hat{k}_1$	Range of Motion $\hat{k}_2$	Derivatized Circumference $\hat{k}_1$	Derivatized Circumference $\hat{k}_2$	Malalignment $\hat{k}_1$	Malalignment $\hat{k}_2$	DJD Score $\hat{k}_1$	DJD Score $\hat{k}_2$
1	0.322	NA					NA	NA
2	0.195	0.418					DEGEN§	NA
3	0.202	NA					0.274	NA
4	0.168	0.272^{NS}					NA	NA
5	0.151	NA					DEGEN§	NA
6	0.210	NA					0.125	NA
7	0.093^{NS}	NA					DEGEN§	NA
8	0.253	0.607					0.128^{NS}	NA
9	0.386	NA	0.436	NA	0.459	NA	0.226	0.399
10	NA	NA	0.472	NA	0.137	NA	0.266	NA
11	NA	NA	0.465	NA	0.324	NA	0.337	NA
12	NA	NA	0.346	NA	0.375	NA	0.174	NA
13	0.225	0.332	0.538	NA	---	---	0.216	0.447
14	0.414	NA	0.516	NA	0.292	0.712	0.267	NA
15	0.345	NA	0.468	NA	0.371	NA	0.250	NA
16	0.470	NA	0.531	NA	0.321	NA	0.282	NA
17	0.320	NA	0.457	NA	0.264	NA	0.370	0.590
Number of hands	128	128	128	128	128	128	128	128

*$\hat{k}_1$ measure perfect agreement and $\hat{k}_2$ agreement within one unit of measurement.

All values are significant ($p < .05$) except where indicated by NS.

NA = not applicable because crude agreement >90%.

§DEGEN = not appropriate because of degenerate contingency table.

RIGHT HAND

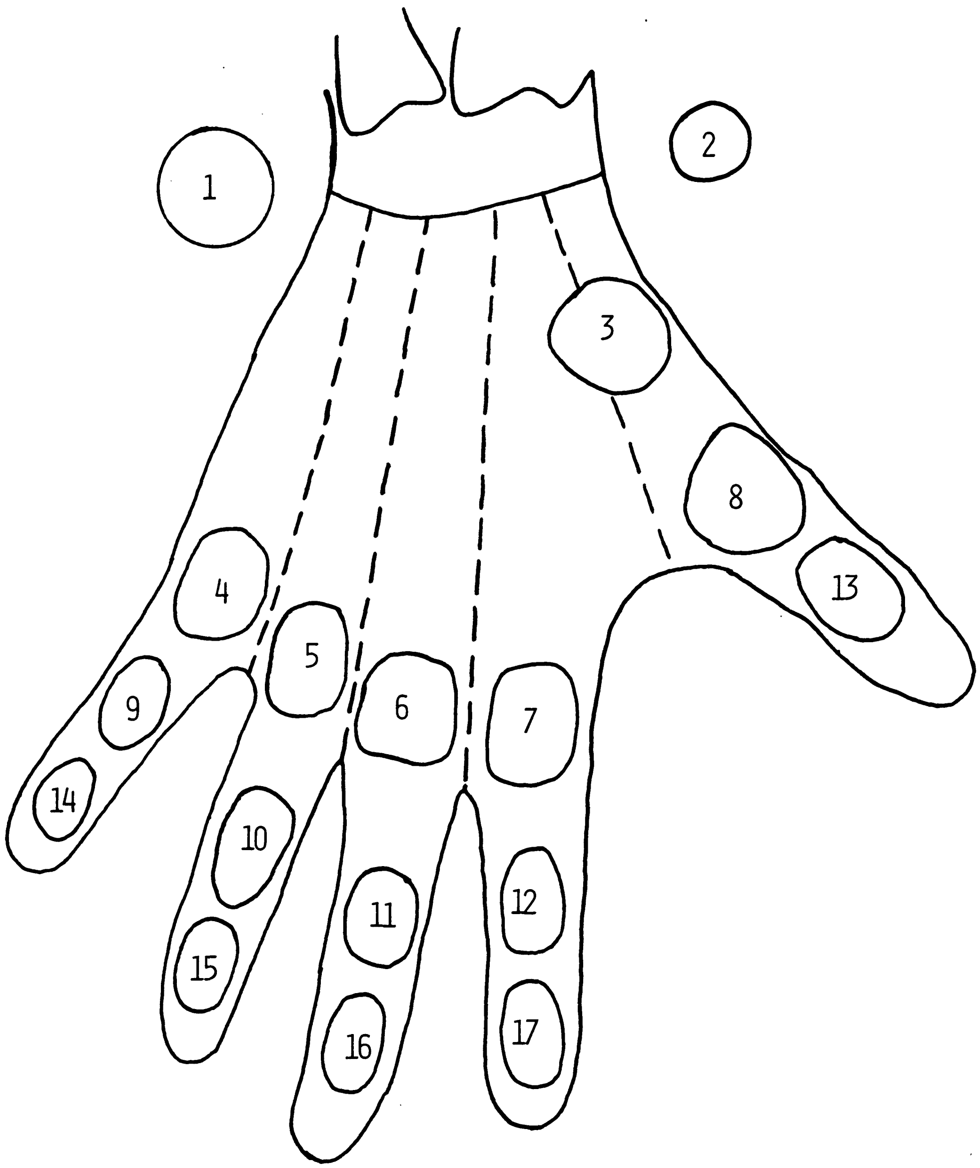
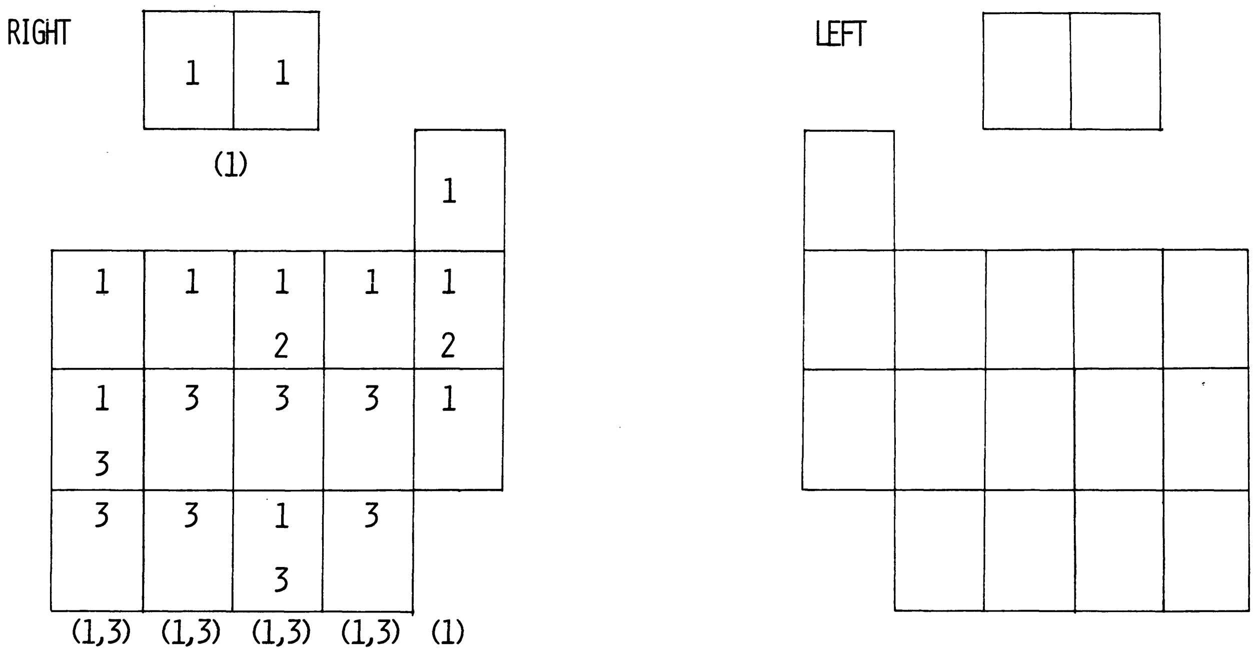


Figure 1

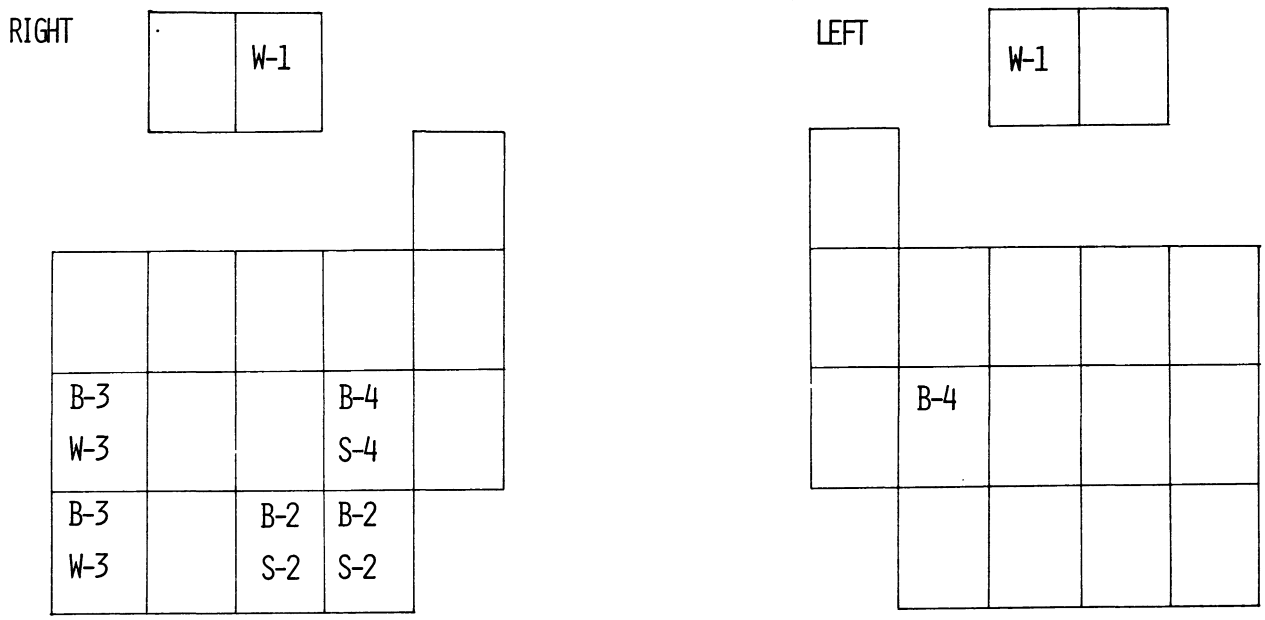
Data collection form and joint numbering code for the right hand. The mirror-image form was used for the left hand. For range of motion data, joint 1 represented the range of deviation at the wrist and 2 represented the range of extension-flexion. For radiographic DJD score joint 1 was the radial-carpal joint, 2 was the inferior radial-ulna joint. Joint 3 represented the first carpo-metacarpal joint throughout. Joints 4-17 are the anatomically corresponding digital joints.



1. Range of Motion
2. D.J.D. Score
3. Derivatized Circumference
4. Malalignment

Figure 2

Differences between right and left hands. The format of this figure is a stylized version of Figure 1. Rejection of the null hypothesis utilizing the multivariate extension of Friedman's chi-square test is indicated by a symbol designating the more impaired side. If the entire digit or wrist is significantly impaired, that is indicated by the designation in parentheses beneath the corresponding region. The data from which the null hypothesis is rejected is indicated numerically: 1 for range of motion, 2 for radiographic DJD score, 3 for derivatized circumference and 4 for radiographic malalignment.



1. Range of Motion
2. D.J.D. Score
3. Derivatized Circumference
4. Malalignment

Figure 3

Task-related differences. The format of this figure is a stylized version of Figure 1. The most consistent and statistically significant structural and functional impairments deduced from the Kruskal-Wallis and MANOVA analyses are noted. When the null hypothesis is rejected, the tasks are ranked and the task(s) with the most impairment is indicated in the box corresponding to the involved anatomical region. The notation for each task is W for winding, B for burling, and S for spinning. The data from which the null hypothesis is rejected is indicated numerically: 1 for range of motion, 2 for radiographic DJD score, 3 for derivatized circumference and 4 for radiographic malalignment.