Reliability of the Gross Motor Function Measure for Children with Osteogenesis Imperfecta

A thesis project submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the Masters of Science Applied in Rehabilitation Science

Joanne Ruck

7111696

School of Physical and Occupational Therapy
McGill University, Montreal, Canada

July 26, 1999

Abstract

RATIONALE: Objective outcome measures in rehabilitation are required to demonstrate the effectiveness of interventions. The Gross Motor Function Measure is a criterion-referenced evaluative tool designed to detect change over time for children diagnosed with cerebral palsy. Reliability of the measure has not been tested with children diagnosed with osteogenesis imperfecta.

PURPOSE: The purpose of this study was to determine the intra and inter-rater reliabilities of the Gross Motor Function Measure (GMFM) for use with children diagnosed with Type I, III or IV osteogenesis imperfecta, by physiotherapists at the Shriners Hospital for Children.

METHODS: One physiotherapist (the author) administered and scored the GMFM on 19 children with osteogenesis imperfecta who were followed at the Shriners Hospital. These children ranged in age from 8 months to 17 years and 11 months. There were 2 children with Type I, 9 with Type III, and 9 with Type IV. The live assessments were videotaped, then viewed and scored independently by 4 pediatric physiotherapists at least 6 weeks later. All therapists had previously passed criterion testing of the measure for children with cerebral palsy. The author also scored the videotapes.

RESULTS: Intraclass correlations (ICC) were used to estimate both the intra and inter-rater reliabilities: model (3,1) for the intra-rater reliability and model(2,1) for the inter-rater reliability. The ICC's for intra-rater reliability of the 5 dimensions and the total score were 0.99. The ICC's for inter-rater reliability were 0.98 for the lying and rolling dimension and 0.99 for the other dimensions and the total score. Kappa was calculated for items that demonstrated more disagreement than the majority. The simple Kappa for items 3,4, and 19 ranged from .396 to 1.010 while the weighted Kappa for items 3 and 19 ranged from .682 to .949.

DISCUSSION: Both the intra and inter-rater ICC's were excellent. Our results are slightly higher than those estimated by the McMaster University GMFM group. Children with osteogenesis imperfecta are all capable of following instructions which facilitates the administration and scoring of the measure. The raters had all passed criterion testing of the measure by the GMFM group and had been trained for use of the measure with osteogenesis genesis prior to our study. The videotape provided consistency for the scoring of the measure. The sample demonstrated heterogeneity as the lowest total score was 8.66% and the highest score was 98.6%. This study provides evidence of the reliability of the GMFM for children with osteogenesis imperfecta when administered and scored by pediatric physiotherapists familiar with the measure.

Acknowledgements

The thesis project is a group effort and I wish to express my gratitude to all those who have assisted me in the completion of this endeavour.

First and foremost, I wish to express my heart felt thanks to Dr. Sharon Wood-Dauphinee for her guidance and support as my graduate advisor. Her knowledge and expertise of health care measures have been instrumental in my development.

I wish to express my gratitude to Dr. James Hanley, my statistical advisor. His excellent teaching style has enabled me to comprehend the statistical concepts required for this reliability study. Thanks are due Mr. Faiz Ahmed who carried out the statistical computations.

Thanks are due Dr. Horatio Plotkin, as external supervisor. His knowledge of osteogenesis imperfecta and willingness to assist me throughout this project are greatly appreciated.

I am also indebted to the staff of the Shriners Hospital physiotherapy department who dedicated countless hours in their participation of the inter-rater reliability section of this study: Louise Loiselle, Rochelle Rein, Huyen Phan and Rita Yap. Louise Toupin provided assistance with the Excel program.

Thanks are due the patients and families who agreed to participate in the videotaping of the GMFM evaluations.

I wish to thank Mr. Denis Alves and Jane Wishart who videotaped the GMFM evaluations, sometimes at a moment's notice.

The Shriners Hospital administration is acknowledged for providing the research setting and leave of absence.

Finally I wish to express my thanks to my son Paul for his clerical computer assistance and my husband for his patience and support during this master's degree.

Table of Contents

List of Tables

List of Appendices

Preface

With the increasing strain on the health care budget, rehabilitation clinicians are being called upon to demonstrate, by means of objective outcome measures, the effectiveness of their interventions.(1) Patients who are followed in physical therapy, should demonstrate clinically meaningful change over time to warrant intensive rehabilitation and utilization of health care resources.

The Shriners Hospital has a large population of children with osteogenesis imperfecta, who are followed medically and receive physical therapy at the hospital or in their local communities. One hundred and fifty children with osteogenesis imperfecta are receiving intravenous disodium pamidronate; a drug, which is demonstrating, increased bone density. [2] Together with the rehabilitation, the children are showing some functional gains. However, the currently used measures of gross motor function, have not been adequately sensitive to detect a clinically significant change in this heterogeneous group of patients. Both physicians and clinicians are convinced that improvements occur but lack the instrument to measure the change.

The Gross Motor Function Measure (GMFM) is an evaluative tool that has been designed for children with cerebral palsy. Because the two populations exhibit some similarities such as developmental delay and hypotonia, it was suggested that the GMFM had the potential as a safe and sensitive measure for children with bone fragility. An intra-rater and a inter-rater reliability study of the GMFM for pediatric physical therapists at the Shriners Hospital has been carried out on a sample of children, diagnosed with Type I, III and or IV osteogenesis imperfecta, who are followed at this institution.

1.0 BACKGROUND

1.1Clinical Features of Osteogenesis Imperfecta

Osteogenesis Imperfecta is a genetic disease of connective tissue, which in 70% of individuals is caused by mutations in one of the two genes (COL1A1 and COL1A2) that encode the type 1 collagen chains. [3-6] The incidence is approximately 6.5 per 100,000 and the prevalence is 1 per 10,000 individuals. [7] The major clinical feature of osteogenesis imperfecta is the bone fragility, which varies from mild to severe and often leads to fractures, progressive skeletal deformities, vertebral deformities and markedly short stature. [8] Studies have shown that the bone mineral density (BMD) in patients with osteogenesis imperfecta is reduced compared with age and gender-matched controls. [9,10] One study found that the mean BMD in the lumbar spine of children with mild osteogenesis imperfecta was only 76.7% of normal. [11] This lower BMD increases the risk for fractures, bowing of long bones and spinal deformity.

The incidence of scoliosis is reported to vary from 30 to 70 %, with congenital type curves increasing rapidly after five years of age. Norimatsu et al [12] found that complications of spinal deformities included respiratory distress, impaired ambulation and diminished activities of daily living. Rowe and Shapiro [13] suggested that weakness of the paraspinal muscles might promote asymmetrical growth of the spine, particularly in young children.

Marked joint hypermobility due to underdevelopment of the ligaments is observed in 70 % of cases. Hypotonia, observed in more severely affected cases, develops from inactivity secondary to disuse from multiple fractures or may be present

due to the underlying connective tissue abnormalities in tendons. [13] Developmental milestones may be delayed or arrested in more severely affected children. [15-17]

Basilar invagination is also a complication of osteogenesis imperfecta, which if left untreated, results in brain stem compression. The clinical manifestations include gradual loss of function of the extremities, parasthesia, ataxia and headache. It has been diagnosed in mild, moderate and severely affected children. [18,19]

1.2 Classification of Osteogenesis Imperfecta

Before 1979, osteogenesis imperfecta was classified according to the time at which fractures first occurred: congenita and tarda. The most widely accepted classification, which was developed by Sillence, Senn and Danks, [20] is based on modes of inheritance, radiological and clinical findings, [11,21,22] Type I is the mildest form of osteogenesis imperfecta, while Type II is lethal in the perinatal period. Type III is the most severe non-lethal form, with frequent fractures, marked deformity and short stature. Type IV includes a heterogeneous group of patients, who do not fit the Type I or Type III profiles.

Table 1: Sillence Classification of OI (Adapted)

Type	Genetic Status	Description
I	Autosomal dominant	Mildest form of OI Mild to moderate bone fragility without deformity Associated with blue sclerae, early hearing loss, easy bruising May have mild to moderate short stature
II	Autosomal dominant or recessive	Perinatal lethal Extreme fragility of connective tissue; multiple in utero fractures; usually intrauterine growth retardation Soft, large cranium Micromelia; long bones crumpled and bowed; ribs beaded
III	Autosomal recessive	Progressive deforming phenotype Severe fragility of bones; usually have in utero fractures Severe osteoporosis Relative macrocephaly with triangular facies Fractures heal with deformity and bowing Associated with white sclerae and extreme short stature, scoliosis
IV	Autosomal dominant	Skeletal fragility and osteoporosis more severe than Type I Associated with bowing of long bones; light sclerae: ± moderate short stature: ± moderate joint hyperextensibility

Gerber LH et al (1998): Effects of withdrawal of bracing in matched pairs of children with osteogenesis imperfecta. Arch Phys Med Rehabil 79: 48.

1.3 Treatment of Children with Osteogenesis Imperfecta

The role of orthopedics in this population is the prevention of deformities and fractures, the correction of deformities and ultimately the improvement of function, in particular ambulation. [23] Intramedullary stabilization of long bones, such as the femur, tibia and humerus, are accepted methods of correcting deformities, which limit ambulation.

Anabolic steroids were used in the past to treat this condition, but without success. [24] A new drug, disodium pamidronate, similar to that used in the treatment of adult osteoporosis, is demonstrating increases in BMD. Subjective improvements of well being, and chronic pain relief have also been described and mobility and ambulation have improved in some children. [2,25]

The goal of physical therapy management of these frequently frail children with multiple fractures, skeletal deformities, hypotonia, joint hypermobility and gross motor delay is maximization of functional independence. [13,22,23] Intervention strategies are based on knowledge of the child's achievement of milestones and focus on improving muscle strength, muscle stabilization of the joints as well as functional ability. [16]

1.4 Gross Motor Function in Osteogenesis Imperfecta

A number of studies have examined the relationship between the achievement of milestones and eventual mobility status.[14,16,17] In infants with Type III or Type IV presentation, developmental milestones are delayed and the order of achievement differs from the sequence of normally expected milestones.[17,28] One study supported the

finding, that the order of achievement of milestones differs from expected milestone dates.[16] Static milestones develop at an earlier stage than dynamic milestones.

Children with osteogenesis imperfecta participating in comprehensive rehabilitation programs that combine physical therapy, lower extremity bracing and orthopedic surgery when indicated have demonstrated high levels of function. Some children gained the ability to ambulate, which they might not have achieved without the comprehensive rehabilitation. [26] The effect of physical therapy alone has yet to be demonstrated. Outcomes of clinically significant change in gross motor function in this population are lacking in the literature.

1.5 Measures of Gross Motor Function used with Osteogenesis Imperfecta

Only a few studies on osteogenesis imperfecta have incorporated standardized measures of gross motor development. Engelbert described a disability profile using the Pediatric Evaluation of Disability Inventory (PEDI). The PEDI is a measure of self-care, care-giver assistance, mobility and social function, which is administered by structured interview and based on parent report. [29] Engelbert reported that Dutch children, under the age of 7.5 years with Type III and IV osteogenesis imperfecta, scored more than 2 standard deviations below the median in the mobility domain.[28]

Bleakney and Kruse [30] assessed 10 children with both the PEDI and the Peabody Developmental Motor Scale. Nine out of 10 children exhibited significant gross motor delays. Change over time was not described. The Peabody Motor Scale is standardized from birth until 83 months. [31] The limitation of administering this scale to children with moderate bone fragility (Type I or IV) is the amount of risk involved.

Examples of items include: walking on a balance beam 2 inches wide, jumping hurdles, jumping in the air while turning, pushups and skipping. Another inadequacy of this measure for this population is the failure to take into consideration the use of adaptive aids such as orthoses, canes, crutches and walkers. Assessments should be able to reflect improvements over time while including the amount of external support required for gross motor function.

The Alberta Infant Motor Scale (AIMS) and the Movement Assessment of Infants (MAI) have been developed as screening tools only for use with infants. Table I [31-38]. These instruments are not applicable for older children. The Bayley Scale of Infant Development assesses both motor and mental development from birth to 36 months of age, compared to a normative sample.[34] This tool is not valid for school aged children. Even in the children with Type I osteogenesis imperfecta, these tools would be appropriate only for young children.

The Basic Gross Motor Assessment (BGMA) and the Bruinick-Oseretsky Test of Motor Proficiency are instruments which measure mild motor dysfunction in school aged children.[35,36] Since these tests require high levels of gross motor function and the exclusion of walking aids, these measures could not identify small increments of change in moderately and severely affected children with osteogenesis imperfecta. Type III and Type IV often require crutches, canes or walkers to ambulate independently.

The Gross Motor Function Measure (GMFM) was constructed specifically for the purpose of evaluating change in gross motor function in children with developmental disabilities, in particular children with cerebral palsy.[38] Appendices 1 and 2. This instrument consists of 88 items which have been grouped into the following dimensions:

lying and rolling, sitting, crawling and kneeling, standing and running and jumping. Each item is scored on a four point ordinal scale. Each dimension contributes equally to the score.

The validation sample included 111 children with cerebral palsy aged 5 to 60 months, 25 with head injury and 34 non-disabled preschool children. Eighty-eight of the 111 children had spastic type cerebral palsy; 23 had non- spastic type cerebral palsy. Only 2 out of 23 were classified as hypotonic.

The intra-rater and inter-rater reliabilities for repeated administration of the measure were estimated by intra class correlations. (ICC) The inter-rater reliability ranged from 0.87 to 0.99 across the 5 dimensions and 0.99 for the total score. The intra-rater reliability ranged from 0.92 to 0.99 across the dimensions and .99 for the total score. Validity was assessed by correlation of change over a period of 4 to 6 months on the GMFM with observer judgement for parents (r=0.54), therapists (r=0.65) and blind assessors (r=0.82) respectively. [38,39]

In another study, the inter-rater reliability of the total GMFM for children with Down syndrome was greater than 0.90 ICC. However, the lying, rolling, crawling, and kneeling dimensions showed more variability than the other dimensions. (0.73 and 0.88 respectively). [40]

Responsiveness of the GMFM was demonstrated in three studies i.e. post rhizotomy, following a fitness program for children with cerebral palsy and post intensive physical therapy. [41-43]

To summarize, the clinical features of osteogenesis imperfecta are well described in the literature and the Sillence classification is known to clinicians familiar with the

condition. The treatment of osteogenesis imperfecta until recently has included orthopaedic surgery, bracing and physiotherapy. Until recently, drug therapy has been unsuccessful. Disodium pamidronate is now demonstrating increased bone density. There are subjective reductions of pain and improved mobility and ambulation in some children. The effects of physiotherapy alone remain to be proven. Outcomes of clinically significant change in gross motor function in this population are lacking in the literature.

A few studies have incorporated standardized measures of gross motor function, but there are limitations of the tools used. A disease-specific tool does not exist to date. The GMFM was designed to detect significant change in gross motor function in children with developmental disabilities in particular cerebral palsy. It has never been tested in a population of children with osteogenesis imperfecta many most of whom exhibit hypotonia and weakness.

2.0 RATIONALE AND OBJECTIVES

2.1 Rationale

Rothstein (1985) stated "that when evaluating measures for clinical use, it is important to consider population-specific reliability for the particular group being measured and for the type of people administering the measurements."[38] In agreement, Streiner and Norman stated that the reliability of a measure is intimately linked to the specific population to which one wants to apply the measure. It cannot be generalized to other populations without being tested. [44]

As this measure was developed for children with cerebral palsy, it has only been tested for conditions, which are primarily neurological in nature. Children with Type III or Type IV osteogenesis imperfecta, particularly as infants are hypotonic and exhibit significant motor delays as do children with cerebral palsy. [13,15-17] None the less the gross motor abilities of children with this condition who maintain an upright position with support are still below the norm. [28,30] Intra and interrater reliability of this measure have never been tested on a population that is primarily orthopedic in nature. A large percentage of individuals with osteogenesis imperfecta present with marked bowed humerii, femora and tibiae. Scoring items which require full hip and knee flexion (items 4-5) as well as complete elbow extension (12-13) may demonstrate poor reliability. These children experience considerable pain from microfractures undetectable by radiograph. Task performance will be reduced by this discomfort and, therefore, reliability may be reduced.

One of the advantages of use of the GMFM for this population is the inclusion of orthoses and walking aids in the calculation of the score. Every dimension may be scored

with or without aids. Progress over time can be detected, as the child requires less bracing and support for ambulation to accomplish the same tasks.

The GMFM also involves less risk than some of the norm-referenced measures of gross motor function for these children with bone fragility and hypermobile joints. Hopping and jumping could be potentially dangerous, however partial scores are given for initiation of the movement; the risk is thus reduced. In the population with cerebral palsy, this instrument is used for infants, children and adolescents, and young adults. The long term goal is that the GMFM will measure change over time in moderately and severely affected children with osteogenesis imperfecta, from infancy to 18 years of age.

2.2 Objectives

The first objective of this study was to evaluate the inter-rater reliability of the dimensions and total score of the GMFM when administered to children (infancy to 18 years of age) diagnosed with Type I, III or IV osteogenesis imperfecta. Inter-rater reliability determines the extent to which consistent scores are obtained by repeated measures of the same patient. For example, in the clinical setting, several therapists would simultaneously rate the same patient. In our study, the GMFM would be considered reliable if the intraclass coefficients (ICC) were 0.90 for the total score and 0.85 for each of the five dimensions.

The second objective was to evaluate the intra-rater reliability of the measure when comparing a live observation against a videotape of the same evaluation. True intra-rater reliability determines the consistency of repeated measures by the same rater over a short period of time. A high reliability could indicate the most optimistic upper

limit of using this medium for clinical research purposes, since repeated measures are not always practical or feasible in the hospital setting.

3.0 METHODS

3.1 Study Population: raters

The population of raters included all 5 physiotherapists employed at the Shriners Hospital for Children in Montreal. Due to the small size of the physiotherapy department, the author was included in the group of raters. The author has 21 years of clinical experience in pediatrics. The remaining therapists' pediatric physiotherapy experience ranged from 5 to 8.5 years, the mean being 6.75 years. Their clinical experience as physiotherapists ranged from 6.5 to 8.5 years, the mean being 7.5 years. Four of the five therapists graduated from McGill University, the other from the University of Montreal. All the therapists had administered and scored the GMFM with children diagnosed with cerebral palsy.

3.1 Prior Training for Cerebral Palsy

Three of the five therapists, including the author, had been trained by the McMaster University GMFM Group and were experienced in the administration and scoring of the measure for children with cerebral palsy. To receive certification, we attended at a one-day training workshop, which included videotapes of children with varying severity and types of cerebral palsy. A level of agreement with the criterion videotape at kappa greater than or equal to 0.80 was required for certification of the scoring of the measure. This workshop assessed the ability of participants to view and score a sample of items from the GMFM. The remaining 2 therapists were instructed by the author in the administration and scoring of the measure for use with children diagnosed with cerebral palsy.

The entire group was criterion tested last October (1998) for use of the GMFM with children diagnosed with cerebral palsy. A videotape of several items, provided again by the McMaster University GMFM group, was scored. The level of agreement per therapist with the criterion videotape ranged from a 0.86 to a 0.97 weighted kappa. All therapists exceeded the required 0.80 weighted kappa level of agreement.

3.12 Prior Training for Osteogenesis Imperfecta

Before commencement of this study, the four therapists attended another training session on the administration and scoring of the instrument for children with osteogenesis imperfecta. They all practiced administering the GMFM with children with this diagnosis imperfecta. The author developed a sheet of instructions to assist in the scoring of the GMFM for either cerebral palsy or osteogenesis imperfecta. Videotapes of GMFM evaluations of two children, with Types III and IV were individually scored by each member. The group then discussed the videotapes. These two children were not included in the sample population.

3.2 Study Population: patients

The target population included infants, children and adolescents until the age of 18 years diagnosed with Type I, III or IV osteogenesis imperfecta. Any child with osteogenesis imperfecta, who had a fracture, confirmed by radiograph, within 6 weeks prior to the evaluation was excluded. Any child with osteogenesis imperfecta, who was in traction or immobilized in splint, back slab or cast for a recent fracture or post surgically,

at the time of the evaluation, was also excluded. Administration of the GMFM could not be performed under these conditions.

The available population included the 150 patients with Type I, III and IV osteogenesis imperfecta that were followed at the Shriners Hospital in Montreal and were part of the cyclical pamidronate protocol. The Shriners Hospital is an elective pediatric orthopedic hospital, which also specializes in the research of genetic and metabolic diseases in children, including osteogenesis imperfecta. The catchment area for patients includes all of Canada, the New England States and selected patients from around the globe. A growing number of patients from across the U.S.A are now being treated with the pamidronate therapy at this hospital. Osteogenesis imperfecta patients from outside the province tend to have more serious clinical presentations. However, the available population is still representative of the spectrum of clinical manifestations of the condition.

A sample of 19 children was selected according to age and severity of presentation of the condition. (Type I, III, and IV). Based on clinical experience, these children appear to be representative of the spectrum of patients treated at the hospital. Appropriate institutional consent was obtained from each parent prior to the videotaped sessions. Appendix 3. The clinical profiles described in Table 2 illustrate the heterogeneity of the sample. All the children were on the cyclical pamidronate protocol. There were 9 boys and 10 girls. They ranged in age from 8 months to 17 years and 11 months. The mean age was 7.89 years and the median was 6 years. The patient group included 2 children with Type I, 9 children with Type III and 8 children with Type IV osteogenesis imperfecta.

Table 2. Sociodemographic and clinical characteristics of the subjects (n=19)

Subjects	Type	Gender	Age at evaluation
1	1	M	4 yrs. 2 mths
2	1	M	10 yrs. 4 mths.
3	3	M	8 mths.
4	3	F	1 yr. 8 mths.
5	3	F	4 yrs. 6 mths.
6	3	F	5 yrs. 7 mths.
7	3	F	6 yrs.
8	3	F	6 yrs.
9	3	F	6 yrs.
10	3	M	11 yrs. 5 mths.
11	3	F	15 yrs. 7 mths.
12	4	M	5 yrs. 1 mth.
13	4	F	5 yrs. 6 mths.
14	4	F	8 yrs.
15	4	M	8 yrs. 5 mths.
16	4	M	8 yrs.11mths.
17	4	F	10 yrs. 9 mths.
18	4	M	13yrs. 4mths.
19	4	M	17 yrs.11mths.

3.3 Data Collection

The author administered the GMFM and scored the evaluations from the live observations between June 1, 1998 and March 29,1999. The evaluations were videotaped by one of two audiovisual technicians.

A number of studies have used patient videotapes to permit multiple raters to observe the same performance. [1,39,43] According to Gross and Conrad, videotaping permits less biased estimates of reliability. It also facilitates scheduling of patient evaluations when the inter-rater reliability of several raters is involved. [45] Clearly, organizing several therapists to evaluate one patient at the same time is not feasible in

most clinical settings including ours. In this study, the child was videotaped from angles that best permitted complete viewing of the specific task. Occasionally, the best view was not obtained which made scoring from the video more difficult. This is one of the disadvantages of videotaping but it was consistent across therapists.

A minimum of six (6) weeks later, from February 1 to May 20, 1999 the videotapes were scored individually by the four (4) other physiotherapists employed at the Shriners Hospital. The author also scored the evaluations again, the second time from the videotape. This time frame eliminated the possibility of memory of the first score. The therapists were asked not to discuss their scores. Over the several week period of scoring of the videotapes, the therapists realized that some items were more difficult to rate than others and the author referred them back to the GMFM manual.

3.4 Data Recording

The therapists tallied the data, including the score for each dimension and total score. The best of three tries was used. All the patients were given the opportunity to attempt all items of the GMFM evaluation that were deemed safe. However, any tasks, which were perceived by the family/ child or physiotherapist to put the child at risk for fracture, were not attempted. Each of the 88 items is scored on a 4-point scale. Values are assigned from 0 to 3, depending upon the percentage of acquisition:

0- does not initiate

1- initiates (less than 10% of the task)

2- partially completes (10 to less than 100% of the task)

3- completes

Each dimension has a different number of items, therefore a different maximum score.

Dimension	**Number of items**	**Maximum score**
A= lying and rolling	17	51
B= sitting	20	60
C= crawling and kneeling	14	42
D= standing	13	39
E= Walking, running and jumping	24	72

The raw score in each dimension is converted into a percentage of the maximum per dimension. In this study, the percentage was carried to the first decimal point. (I.e. 30/51= 58.9%) Each dimension is equally weighted and the total score is calculated by summing the percentages of each dimension and dividing by 5. In this study the total was calculated to the second decimal point. The author entered the percentage of acquisition in each dimension and of the total score on a spread sheet program (Excel)

3.5 Data Analysis

The intra-class correlation coefficient (ICC) was calculated as it is an appropriate measure of agreement for ordinal data that meets the assumption for item summation. The ICC is the ratio of the variance between patients over the variance between patients plus the variance between raters plus error. Perfect concordance without any variance in scores will yield a value of one. ICC was the measurement of agreement used to evaluate the reliability of the GMFM for children with cerebral palsy.[38] In their study, An ICC of 0.75 was considered acceptable for all reliability coefficients.

Reliability coefficients of 0.80 and greater are considered high. However, when a measure is to be used for clinical decision-making for an individual patient, more stringent criteria are recommended.[46,47]. An ICC of 0.90 for a total score is generally accepted to be the minimum required for clinical decision-making.

3.51 Inter-rater reliability

The inter-rater reliability was calculated with the ICC as derived by a 2 way random effects analysis of variance as described by Shrout and Fleiss. Their Model 2,1 is recommended for use in inter-rater reliability studies, where all subjects are evaluated by each of a number of raters, who are considered representative of a larger population of similar raters.[48] It was our opinion that the physiotherapists at the Shriners Hospital were representative of all physiotherapists with some experience working in pediatric settings and who are familiar with the measure.

$$(1)\ ICC = \sigma^2_{patients} / \sigma^2_{patients} + \sigma^2_{raters} + \sigma^2_{error}$$

The unbiased estimators of the three components of variance are:

$$S^2_{patients} = MS_{patients} - MS_{error} / no.\ patients$$

$$S^2_{raters} = MS_{raters} - MS_{error} / no.\ patients$$

$$S^2_{error} = MS_{error} / no.patients$$

MS= mean square

Each of the last three last formulae substitutes back to formula (1). After simplification, this is equivalent to calculating the ICC directly as:

$$= \frac{no.patients(MSpatients - MSerror)}{no.\ patients(MSpatients) + no.\ raters(Msraters) + [no.patients(no.raters) - no.patients - no.raters]MSerror}$$

An approximate one sided 95% Confidence Interval (random raters) for the ICC was calculated using equation 1.68 according to Fleiss.[49]

3.52 Intra-rater reliability

For intra-rater reliability, Model 3 derived from a one-way random effect analysis of variance is recommended. For a single rating the formula is:

$$ICC(3,1)= MS_{patients} - MS_{error} / MS_{patients} + MS_{error}$$

It is the estimate of the ratio of the difference of patient variance and the error variance over the sum of the patient and error variance. An estimate for the ICC intra-rater reliability was calculated using equation 1.2 of Fleiss.[49]

3.53 Kappa

Cohen's Kappa is a chance-corrected measure of agreement which describes inter-rater agreement beyond what is expected by chance alone, as reflected by crude agreement.[50]. Perfect agreement is indicated by a value of 1 for Kappa and 0 for chance agreement alone. According to Landis and Koch, values greater than 0.75 are usually considered to represent excellent agreement between raters. Values between 0.40 and 0.75 represent moderate agreement and those below 0.40 represent poor agreement.[51]

Weighted Kappa is an estimate of percentage agreement, correlated by chance and based on weights reflecting the degree of the amount of disagreement. [52] In our study, we estimated both simple Kappa and weighted Kappa for those items demonstrating higher levels of disagreement than the majority.

4.0 RESULTS

Figures 1 through 6 illustrate the degree of agreement among raters for the entire sample. The heterogeneity of the sample is also illustrated by the placement of the groupings in the graphs. None of the total or individual dimension patient scores overlap with each other.

Figure 1 indicates the patients score for dimension A, lying and rolling. The first 5 recordings are the scores given by raters 1 through 5 for the first patient in Table 2. The next five scores are for the second patient and in the same order of raters. (etc.) Figure 2 indicated the results of Dimension B, etc.

The inter-rater and intra-rater reliability results are shown in Table 3. We established a priori that the ICC for the total score should be at least 0.90 and 0.85 for each dimension.

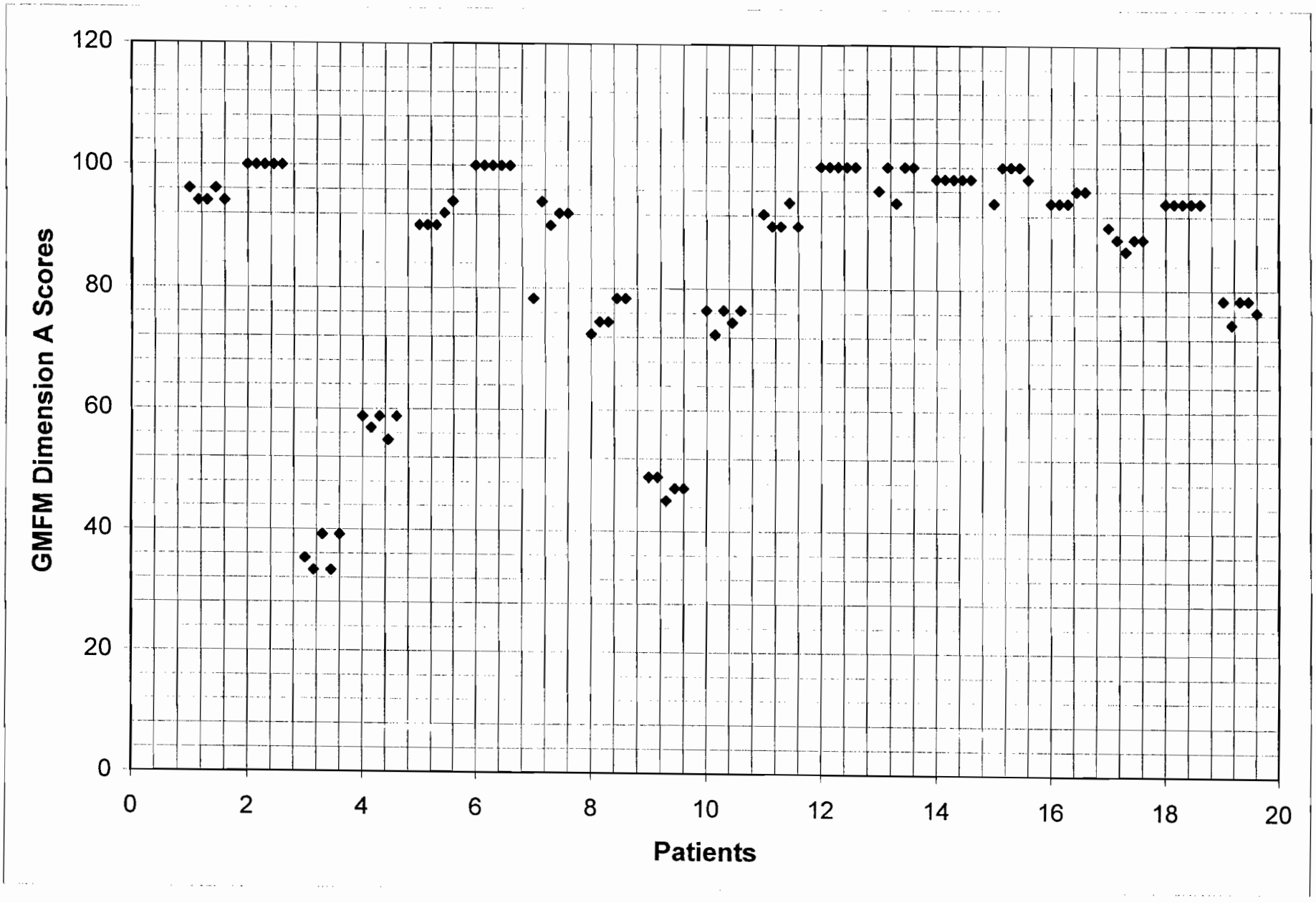

Figure 1: GMFM Dimension A, lying and rolling scores given by the five physiotherapists for the nineteen patients. The scores are grouped in the same order as the patient list in Table 2. The raters are illustrated in the same order from one to five.

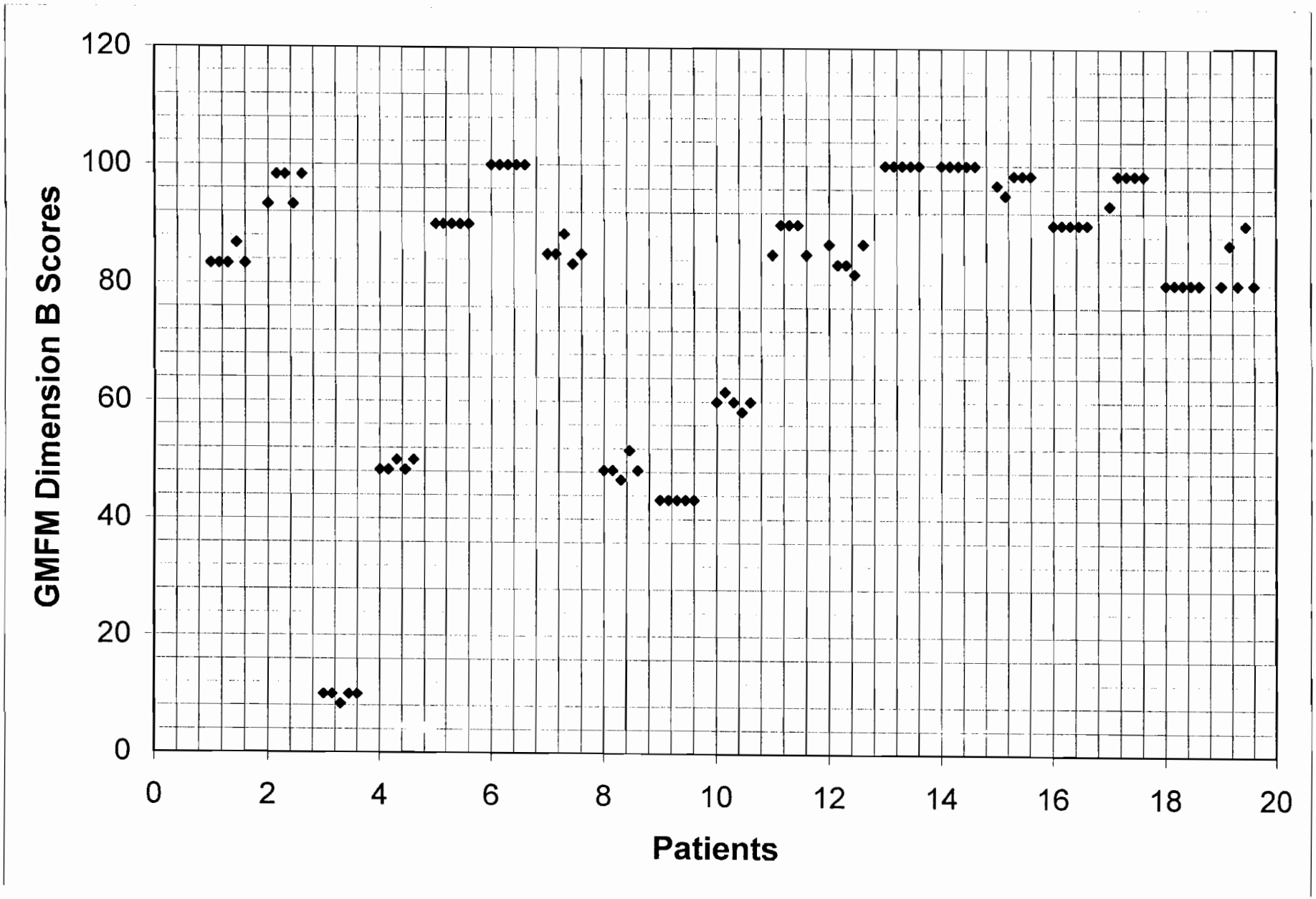

Figure 2: GMFM Dimension B, sitting scores given by the five physiotherapists for the nineteen patients. The scores are grouped in the same order as the patient list in Table 2. The raters are illustrated in the same order from one to five.

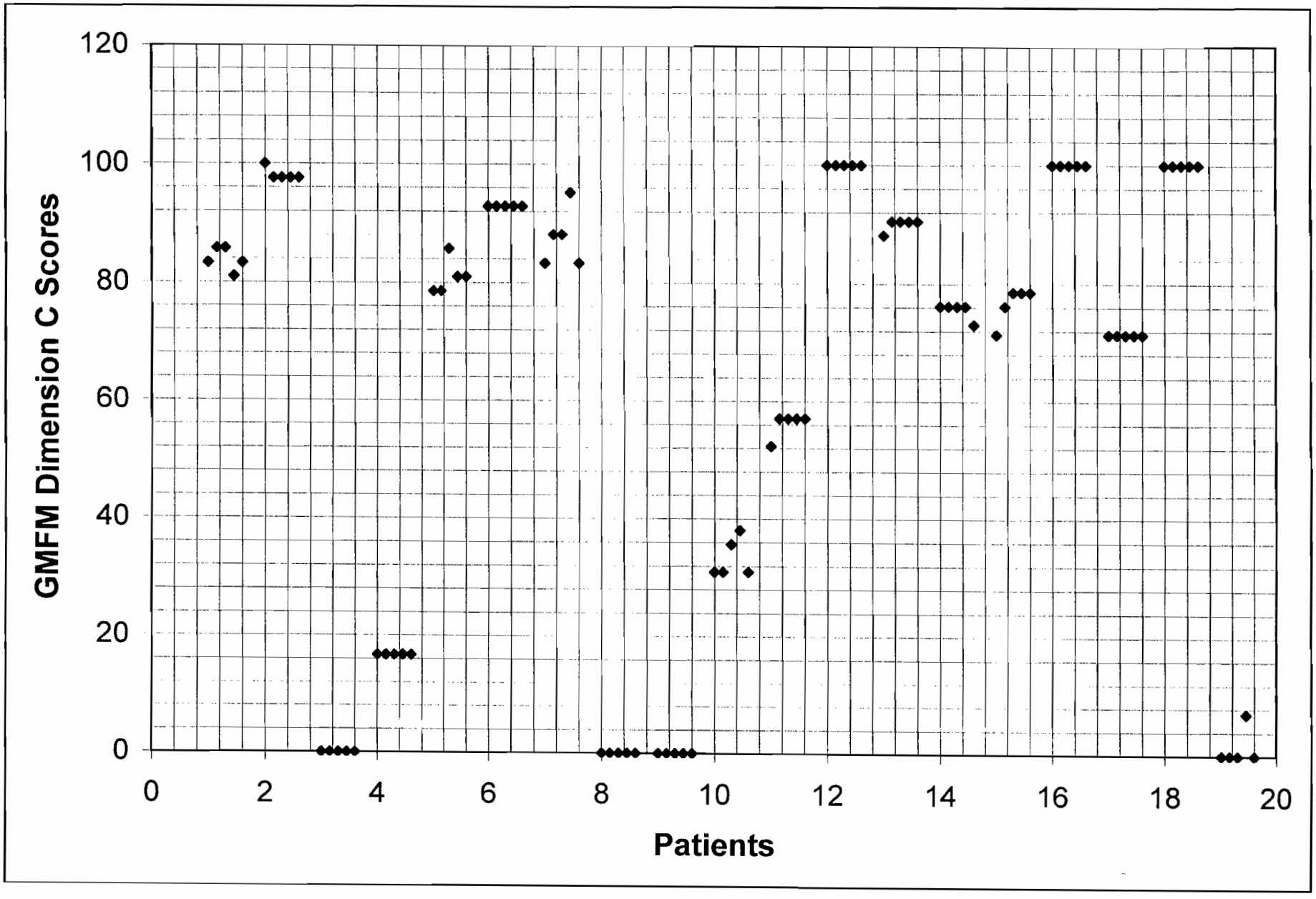

Figure 3: GMFM Dimension C, crawl kneel scores given by the five physiotherapists for the nineteen patients. The scores are grouped in the same order as the patient list in Table 2. The raters are illustrated in the same order from one to five.

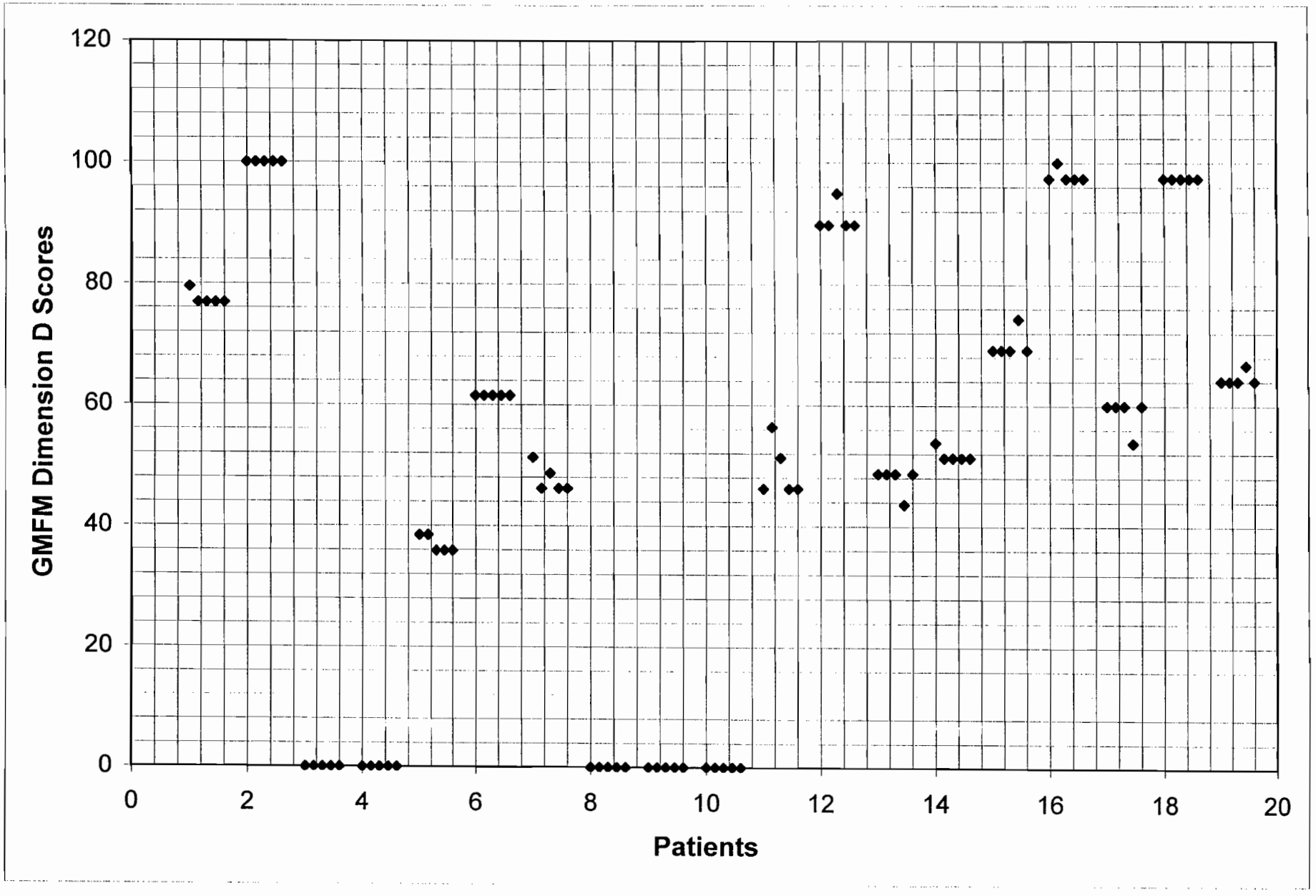

Figure 4: GMFM Dimension D, standing scores given by the five physiotherapists for the nineteen patients. The scores are grouped in the same order as the patient list in Table 2. The raters are illustrated in the same order from one to five.

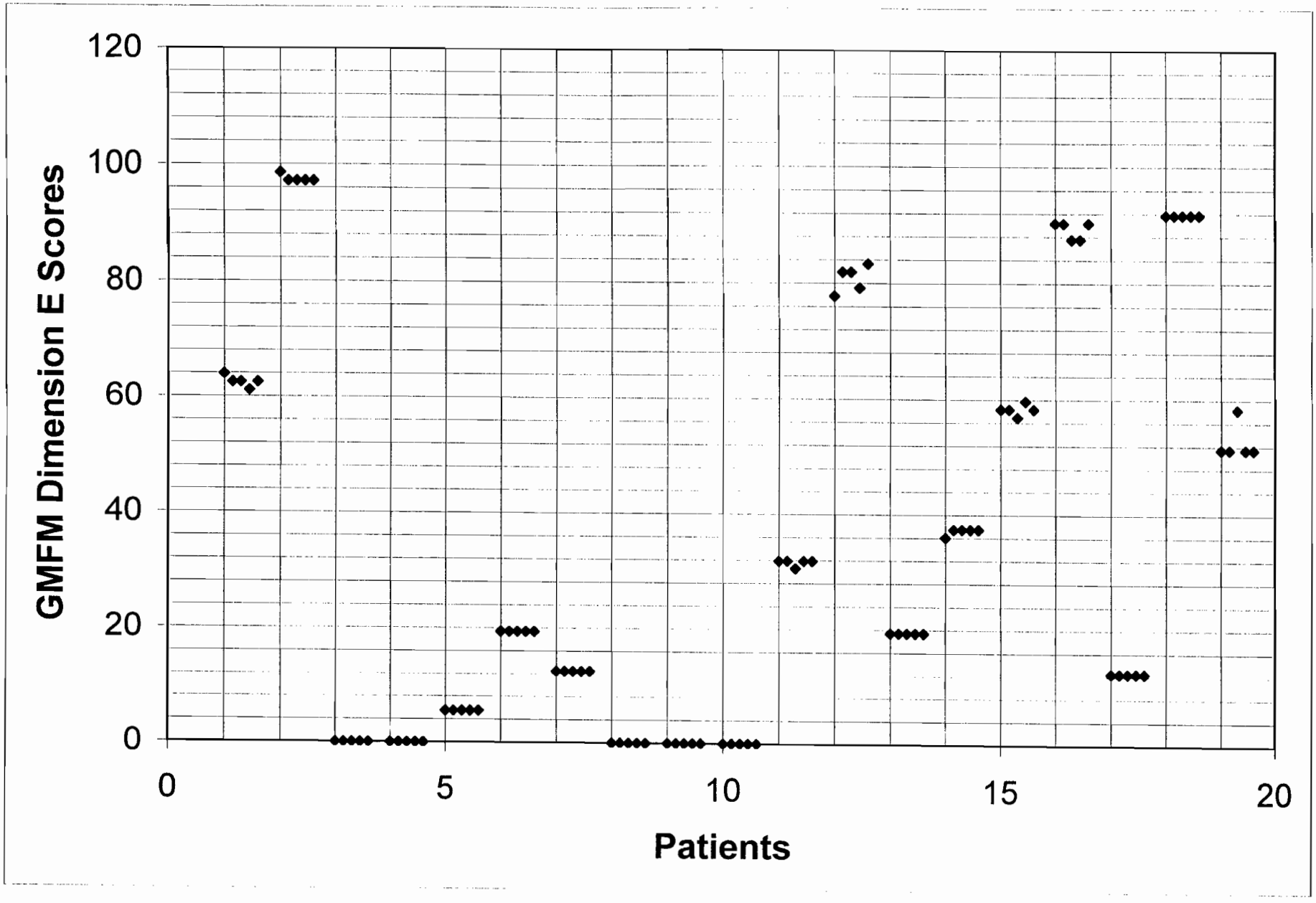

Figure 5: GMFM Dimension E, walk, run jump scores given by the five physiotherapists for the nineteen patients. The scores are grouped in the same order as the patient list in Table 2. The raters are illustrated in the same order from one to five.

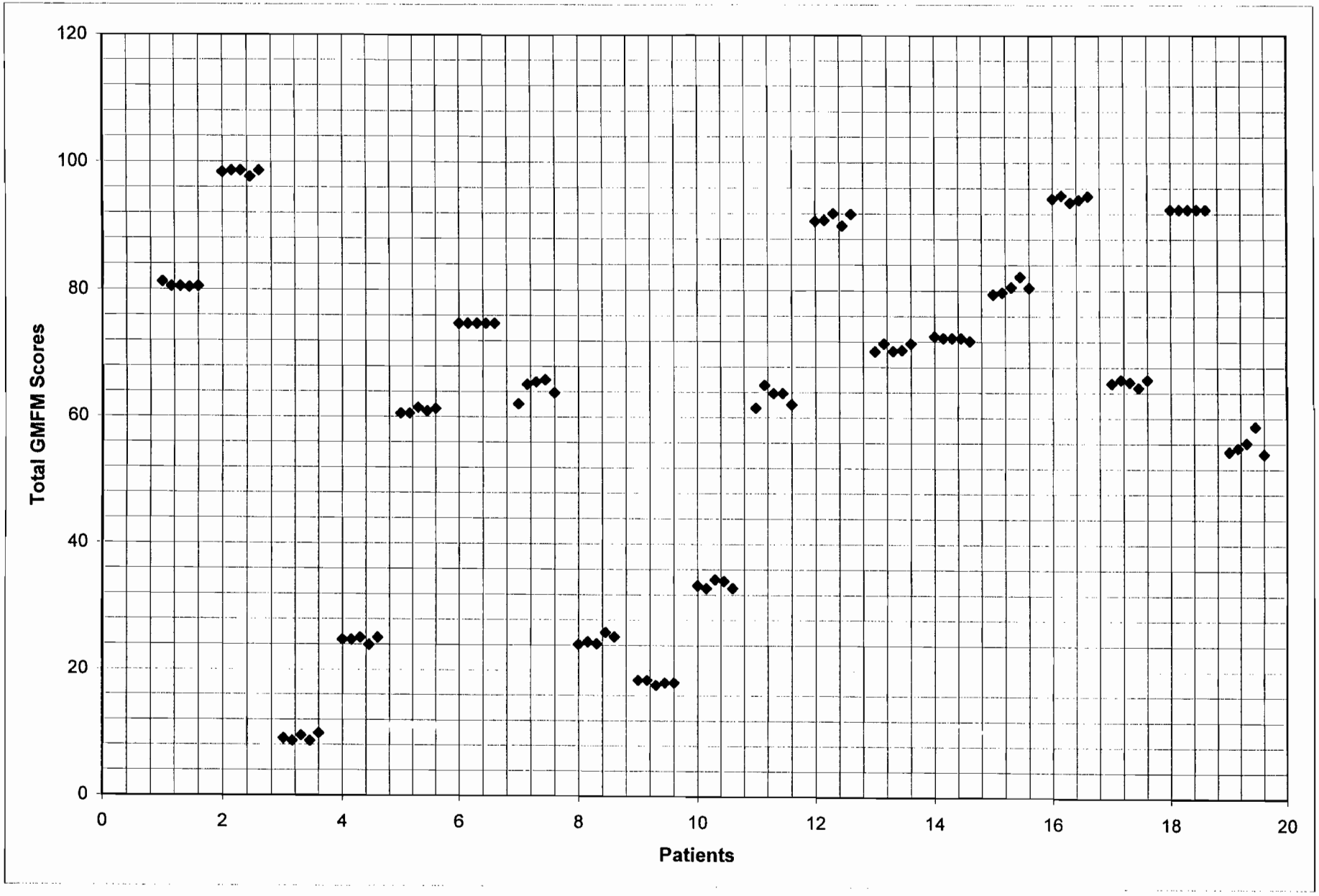

Figure 6: GMFM total scores given by the five physiotherapists for the nineteen patients. The scores are grouped in the same order as the patient list in Table 2. The raters are illustrated in the same order from one to five.

Table 3. ***Reliability of the Dimension Percent Scores and the Total Percent Scores of the Gross Motor Function Measure : Intra-Class Correlation Coefficients (ICC 2,1) and lower Confidence Limits.***

	Intra-rater		Inter-rater	
	19 patients, twice; lower 95%confidence limit		19 patients, 5 raters;lower 95% confidence limit	
Dimension	**ICC**	**confidence limit**	**ICC**	**confidence limit**
Lying and rolling	0.99	0.99	0.98	0.97
Sitting	0.99	0.99	0.99	0.99
Crawling and kneeling	0.99	0.99	0.99	0.99
Standing	0.99	0.99	0.99	0.99
Walk, run & jump	0.99	0.99	0.99	0.99
Total score	0.99	0.99	0.99	0.99

In this population of children with osteogenesis imperfecta, the individual dimension and total score inter-rater reliabilities were greater than 0.98.The lower confidence intervals of the ICC's for the total percent scores varied from 0.97 to 0.99.

Simple Kappa , estimated by pairs of raters for items 3, 4 and 19 ranged from 0.552 to 0.913. One pair of raters, the third and fifth, scored a perfect Kappa for item 4. The 95% confidence limits were from 0.396 to 1.010. Appendix 4. Weighted Kappa, could be calculated for only items 3 and 19 and ranged from 0.682 to 0.949. The confidence limits were from 0.396 to 1.010. Appendix 5. Weighted Kappa was not computed for item 4 as there was insufficient variance of the scores between pairs of raters.

5.0 DISCUSSION

5.1 Overview

An intra and inter-rater reliability study of scoring of the Gross Motor Function Measure for children with Types I,III or IV osteogenesis imperfecta was carried out by physiotherapists at the Shriners Hospital for Children in Montreal.

These results indicated excellent intra and inter-rater reliability when pediatric physiotherapists score the GMFM for children diagnosed with osteogenesis imperfecta. The inter-rater reliability estimated by intraclass correlation (ICC) ranged from 0.98 to 0.99 across the 5 dimensions and the total score. The lower confidence limits ranged from 0.97 to 0.99. The lowest ICC and confidence limits were noted in the first dimension: lying and rolling. The intra-rater reliabilities estimated with a live evaluation compared to videotape of the evaluation were consistently 0.99 for all dimensions and the total score.

Despite these exceptionally high ICC values, it was apparent that a few items demonstrated more disagreement than the majority: items 3,4 and 19. While weighted Kappa was planned for all three items, item 4 did not have sufficient disagreement to calculate a weighted Kappa. Simple Kappa, computed for items 3,4 and 19 ranged from 0.552 to 0.913. The confidence limits were 0.396 to 1.010. One pair of raters had perfect agreement for item 4. Weighted Kappa computed for items 3 and 19 ranged from 0.682 to 0.949. The confidence limits were 0.396 to 1.010. Weighted Kappa was not computed for item 4 as there was insufficient variance of the scores between pairs of raters to warrant computation. In summary, despite disagreement on a few items, both intra and inter-rater reliability estimates for the GMFM when used with children with osteogenesis imperfecta were exceptionally high.

5.2 Interpretation of the Findings

Our findings were compared with other reliability studies of the GMFM found in the scientific literature. The reliability of the GMFM has only been estimated in two populations: cerebral palsy and Down Syndrome. Three studies are described in the literature.

Our results with a sample of children diagnosed with osteogenesis imperfecta are somewhat higher than those estimated for children with cerebral palsy. While, the intra and inter-rater reliability results obtained by Russell et al were very similar to those obtained in this study, there was one exception, the lying and rolling dimension.[38] Russell's group estimated the lying and rolling dimension as 0.87 ICC for children with cerebral palsy while our results with osteogenesis imperfecta were 0.98. Their reliability for the sitting dimension was 0.92 while ours was 0.99.

There are several possible explanations as to why we achieved such high estimates of reliability. Specifically the inter-rater reliabilities of dimensions A and B may have been higher than with the children with cerebral palsy. Children with osteogenesis imperfecta have normal intelligence and motor planning skills. They are capable of following instructions. In many cases such is not the case when testing children with cerebral palsy. Scoring items in these 2 dimensions may be facilitated by the cooperation of the children with osteogenesis imperfecta. The highest levels of function in children with spastic quadriplegia and cognitive impairments are support sitting and perhaps commando crawling. The inter-rater reliability may be higher for osteogeneis imperfecta than that for the children with cerebral palsy when cognition and motor planning is less than optimal.

Children with osteogenesis imperfecta are rarely impeded by spasticity but rather

affected by hypotonia. The degree of spasticity can alter gross motor ability e.g. ability to sit in long sitting and the amount of upper extremity support for sitting, kneeling and standing. It may be affected by several factors including the child's emotions. When excited or stressed, the degree of tone increases and the level of function is hampered. On repeated assessments, the level of spasticity may change and thus affect the gross motor score. This is not the case for children with osteogenesis imperfecta. Repeated measures of the GMFM in a population of children with cerebral palsy may have resulted in lower scores for both the inter and intra-rater reliability, as the child's gross motor capacity may have altered dependent upon the degree of spasticity.

Bjornson, Graubert et al evaluated the validity of the GMFM for 37 children diagnosed with spastic diplegia who participated in a randomized clinical trial that addressed the efficacy of selective dorsal rhizotomy.[54] As part of their study, they estimated the inter-rater reliability of the six evaluators and the lead physiotherapist. Inter-rater reliability was monitored quarterly from videotapes using the lead physiotherapist responsible for training and supervision as gold standard. They maintained more than a 0.90 point by point agreement. The ICC'S ranged from 0.80 to 1.00. Information regarding dimension scores is not available in the literature nor is the amount of rater training in the administration and scoring of the measure. In our study, all the physiotherapists passed criterion testing of the measure for cerebral palsy.

Russell et al from the McMaster University Neurodevelopmental Clinical Research Unit, conducted a validity study of the GMFM for children with Down syndrome. [40] Two pediatric physiotherapist raters assessed a subsample of 22 children on two occasions separated by a maximum of two weeks. The assessor and observer

roles were determined randomly. The inter-rater reliabilities measured by ICC's for the individual dimensions were estimated at: 0.73 for lying and rolling, 0.97 for sitting, 0.88 for crawling and kneeling, 0.98 for standing, 0.96 for walking, running and jumping. and 0.96 for the total score. The test-retest ICC's were: 0.62 for lying and rolling, 0.96 for sitting , 0.83 for crawling and kneeling, 0.98 for standing , 0.95 for walking, running and jumping and 0.95 for the total score. While children with Down syndrome and osteogenesis imperfecta are both hypotonic, their cognitive abilities are dissimilar. As stated by the authors of the study on Down syndrome, children with this condition who have progressed developmentally beyond a certain dimension, resist performing lower level skills as required by the GMFM. In addition, their ability to follow instructions may be limited by their diminished cognitive capacity. This is not the case for children with osteogenesis imperfecta as their level of comprehension is within normal limits. These reasons may account for the discrepancy in the inter-rater reliability. If the children with Down syndrome did not perform consistently during the two GMFM tests then their test-retest scores would be lower than our intra-rater reliability ICC's. As well, evaluating children with limited cooperation may be more difficult than evaluating children who are cooperative. As a result, the inter-rater reliabilities may be lower in dimensions A and C, lying and rolling and crawling and kneeling since the children would have progressed to the walking stage.

While our results are similar to those of other reliability studies of the GMFM, they exceed the others due to the nature of the population tested. Children with osteogenesis imperfecta have normal cognitive abilities and not spastic rather hypotonic, both of which facilitate rater reliability.

Another possible reason for the high intra and inter-rater reliability is the mode of administration of the test. Specifically, we used videotaped evaluations. A number of issues were taken into consideration when determining the methodology for estimating the intra and inter-rater reliability of the GMFM for children with osteogenesis imperfecta. Since many of the patients are from out of province or even out of country, access to them for a second evaluation was very difficult. Many patients return to the hospital every four (4) months for a period of 3 days, for the cyclical intravenous pamidronate treatment. Test-retest over a 3-day period is too short a period of time to avoid recall. Young children have short attention spans and are easily distracted, therefore, being assessed in front of many raters could result in an inaccurate gross motor score. It was not feasible to liberate 5 physiotherapists for the 19 evaluations during the working day. Therefore, videotapes were used to film the evaluations. Only one physiotherapist was made available for the taping.

All the physiotherapists viewed the videotape from exactly the same angle. Even when the shot was not taken from the best perspective, it was consistent for all the viewers. In a clinical setting, the therapists would observe the evaluation from slightly different positions in the room. Their eyes might focus on different aspects of the task. The videotaping technique standardized the evaluation, which is not possible in a live setting.

In our study, only one physiotherapist (the author) administered the GMFM In a clinical situation, a variety of therapists would administer the GMFM. It is probable that the inter-rater reliability would have not been as consistent had several examiners been implicated in the study. In addition, since the evaluation occurred only once, varying

degrees of noncompliance did not affect the score and thus increase the variation of the intra-rater reliability. The single evaluation may have resulted in an overestimate of the intra-rater reliability.

The high inter-rater reliability may also be explained by the fact that all the therapists have many years of experience with administration of the GMFM for children with cerebral palsy. Their knowledge of normal gross motor development was considerable and refined. Finally, they received considerable training with this measure for both cerebral palsy and osteogenesis imperfecta. Upon criterion testing for cerebral palsy, they scored between 0.86 and 0.97. All the physiotherapists were also trained for use of the measure for children with osteogenesis imperfecta.

Another possible explanation for the high inter-rater reliability was the heterogeneity of the sample population. The lowest GMFM score was 8.66% and the highest was 98.62%. There was no overlap of scores among the raters. Portnoy and Wilkins state that reliability is based on the proportion of the total observed variance that is attributable to error. [53] Therefore, for a given amount of error variance, as the total variance increases the error component accounts for a smaller portion of it. The greater the range of scores, the smaller is the variance due to error and the higher is the inter-rater reliability. In summary, we had an extremely heterogeneous sample, which may have contributed to the excellent reliability results.

There were several dimensions in which patients were consistently scored a zero. Some children were either too young or lacked the strength and balance to accomplish the tasks of the dimension. Other children did not cooperate with the crawling dimension due to fragility of their upper extremities or marked bowing of their tibias, which made

crawling uncomfortable. The consistency of scoring zero for some dimensions may also have resulted in an exceptionally high estimate of intra and inter-rater reliability.

Despite the very positive results, the physiotherapists did however, have some difficulty scoring 3 items as demonstrated by the simple and weighted Kappa computations. Appendices 4 and 5. Items 3 and 19 had the most disagreement. A possible reason for the disagreement in item 3 (supine: lifts head 45 degrees) is the inability to detect from the videotape the active contraction of the neck flexors. Many children with osteogenesis imperfecta are hypotonic and macrocephalic, therefore active flexion of the neck is difficult. Consequently, they elevate their shoulders and passively lift their heads by pushing themselves up with their arms. Distinguishing between the true neck flexion and the compensatory movements from the videotape was not always accomplished.

The instructions for items 19 and 20 are to roll over to one side from the supine position, then attain sitting. Children with osteogenesis imperfecta frequently sit up in the same manner as described in the previous paragraph. When asked to roll over, they barely roll to one side then sit up. Other children with a history of upper extremity fractures avoid prolonged weight bearing on their arms. Again, they avoid rolling completely to side lying and pushing up from this position. The physiotherapists were uncertain whether the patients sufficiently accomplished the task required by this item. Consequently, there some degree of disagreement as measured by the weighted Kappa. (0.682 to 0.780)

Items 4 and 5 require the child flex their hip and knee through full range. While children with osteogenesis imperfecta rarely have contractures, they often present with femoral and tibial bowing. The therapists demonstrated some disagreement, as it

appeared difficult to observe whether the child had achieved sufficient range to fulfill the requirements of the task.

In summary, excellent results were obtained from an intra and inter-rater reliability study of the GMFM conducted by physiotherapists at the Shriners Hospital for children with osteogenesis imperfecta. Our results are higher than those estimated for children with cerebral palsy or Down syndrome. Children with osteogenesis imperfecta have normal cognitive function and are able to follow instructions. They are not affected by spasticity rather hypotonia. Spasticity, which is enhanced by a variety of circumstances including emotions, can limit the children's ability to assume and maintain positions. Thus, gross motor function may be inconsistent and the reliability reduced. Videotaping, our mode of administration of the test, standardized the evaluations and thus increased the inter-rater reliability. The physiotherapists had all been trained in use of the GMFM and had passed criterion testing prior to the start of the study. (weighted Kappa 0.86-0.97) The heterogeneity of the sample reduced the amount of variance due to error and thus increased the inter-rater reliability. Despite the excellent results there were a few items where there was more rater disagreement than for the majority.

The main limitation of the study is the mode of administration of the test. The videotaped evaluations provided a medium for excellent reliability but did not replicate the clinical setting. The ability to score the GMFM was tested but not the ability to administer the measure. The inter-rater reliability of a live evaluation was not conducted.

5.3 AVENUES FOR FURTHER RESEARCH

Establishing reliability of the GMFM for children with osteogenesis imperfecta was a first step towards using this potentially useful instrument in the measurement of clinically significant change of gross motor function. At the present time, little is known about the potential for change in the more severely affected children, due to the inadequacies of the available measures of gross motor function. A more sensitive measure is required to objectively demonstrate the efficacy of rehabilitative intervention programs, medical interventions such as intramedullary stabilization of long bones [23,55-57] or drug therapy to increase BMD, [2,25] all of which could impact upon the child's motor function. Parents and clinicians would have a measure which reflected the progress (or deterioration) of gross function. The utilization of health care resources could more easily be justified with a sensitive outcome measure of gross motor function.

A responsiveness study of the GMFM for use with children diagnosed with ostegenesis imperfecta, who are undergoing rehabilitation, would be an important future study. This study would determine whether this tool could measure clinically significant change in this heterogeneous group of children.

7.0 CONCLUSION

While the validity of the GMFM for use with children diagnosed with osteogenesis imperfecta has not been addressed in this study, excellent reliability has been demonstrated. Pediatric physiotherapists can be trained to score gross motor function with precision.

The Gross Motor Function Measure has proven to be a reliable and safe measure for children diagnosed with osteogenesis imperfecta. The next step would be to determine the responsiveness of this measure over time in this population.

Despite the limitations, videotaping has been shown to be an effective and practical means to estimate inter-rater reliability.

References

1. Badke M, Di Fabio, Leonard E (1993): Reliability of a Mobility Assessment tool with Application to Neurologically impaired patients: a preliminary report. Physiotherapy Canada 45:1, 15-20.

2. Glorieux F, Bishop N, Plotkin et al (1998) Cyclic administration of pamidronate in children with severe osteogenesis imperfecta.NEJM:339:947-952.

3. Byers PH (1989): Disorders of collagen biosynthesis and structure. in: The Metabolic Basis of Inherited Disease ; Scriver CR, Beaudet AL, Sly WS, Valle D, 6th ed, McGraw-Hill, New York,2805-2842.

4. Byers PH (1990): Brittle bones-fragile molecules: Disorders of collagen gene structure and expression. Trends Genet 6:293-300.

5. Prockop DJ, Constantinou CD, Dombrowski KE, Hojima Y, Kadler KE, Kuivaniemi H, Tromp G, Vogel BE (1989) Type 1 collagen: The gene-protein system that harbors most of the mutations causing osteogenesis imperfecta and probably more common heritable disorders of connective tissue. Am J Med Genet 34:60-67.

6. Kuivaniemi H, Tromp G, Prockop DJ (1991): Mutations in collagen genes. Causes of rare and some common diseases in humans. FASEB J 5:2052-2060.

7. Byers PH , Steiner RD(1992): Osteogenesis imperfecta. Annu Rev Med 43: 269-282.

8. Castells S (1973): New Approaches to Treatment of Osteogenesis Imperfecta. Clinical Orthopaedics and Related Research 93: 239-249.

9. Escalante A, Beardmore TD (1993): Decreased bone mineral density in HLAS.B27 positive members of a family with osteogenesis imperfecta. J Rheum 20:320-324.

10. Kurtz D, Morrish K, Shapiro J (1985): Vertebral bone mineral content in osteogenesis imperfecta. Calcif Tissue Int. 37:14-18.

11. Zionts L, Nash J, Rude R (1995): Bone Mineral Density in Children with Mild Osteogenesis Imperfecta. J Bone Joint Surgery (Br) 77-B143-147.

12. Norimatsu H, Mayuzumi T, Takahashi H (1982): The Development of the Spinal Deformities in Osteogenesis Imperfecta, Clinical Orthopaedics and Related Research 162: 20-25.

13. Rowe D, Shapiro J (1990): Osteogenesis Imperfecta: in Metabolic Bone Diseases and Clinically Related Disorders: Avioli L, Krane S (eds), WB Saunders Philadelphia: 659-701.

14. Vetter U, Pontz B, Zauner E, Brenner E, Spranger J (1992): Osteogenesis Imperfecta: A Clinical Study of the First Ten Years of Life. Calcif Tissue Int 50:36-41.

15. Charnas L, Marini J (1993): Neurologic and Developmental Outcome in Osteogenesis Imperfecta. Proceedings of Fifth International Conference on Osteogenesis Imperfecta. Oxford UK.p.50

16. Engelbert R, Helders P, Keeson W (1995): Intramedullary rodding in Type III osteogenesis imperfecta Effects on neuromotor development in 10 children. Acta Orthop Scand 66(4): 361-364.

17. Daly K, Wisbeach A, Sanpera (1996): The Prognosis For Walking In Osteogenesis Imperfecta . J Bone Joint Surg Vol 78-B (3) 477-480.

18. Sawin P, Menezes (1997): Basilar invagination in osteogenesis imperfecta and related osteochondrodysplasias: medical and surgical management. J Neurosurg 86: 950-960.

19. Charnas L, Hopkins, Koby M (1996): Basilar invagination : frequency, detection, progression, and treatment in children with osteogenesis imperfecta .Proceedings of. Sixth International Conference on Osteogenesis Imperfecta. Netherlands.p.35.

20. Sillence D, Senn A, Danks D. (1979): Genetic heterogeneity in osteogenesis imperfecta AM J Med Gen 16: 101-116.

21. Chevrel G, Meunier PJ. (1997): Are drugs helpful in adults with osteogenesis imperfecta? Rev Rhum.[Eng. Ed.] 64(5) : 283-286.

22. Byers P. (1993): Osteogenesis Imperfecta. in Connective Tissue and Its Heritable Disorders Wiley Liss Inc. 317-350.

23. Porat S, Heller E, Seidman D (1991): Functional Results of Operation in Osteogenesis Imperfecta ; Elongating and Nonelongating Rods. J Ped Ortho 10: 200-203.

24. Catell HS, Clayton B (1968): Failure of Anabolic Steroids in the Therapy of Osteogenesis Imperfecta. J Bone Jt. Surg 50-A: 123-141.

25. Soderhall S, Astrom E, Skoog L (1993): Improvement of pain and life quality during APD treatment or a girl with osteogenesis imperfecta type III. Proceedings of the Fifth International Conference on Osteogenesis Imperfecta, Oxford, England.p.78.

26. Gerber L, Binder H, Weintrob J (1990): Rehabilitation of Children and Infants with Osteogenesis Imperfecta A Program for Ambulation. Clinical Orthopaedics and Related Research 251:254-262.

27. Binder H, Conway A, Nason (1993): Comprehensive Rebilitation of the Child with Osteogenesis Imperfecta. AM J Med Gen 45: 265-269.

28. Engelbert R, Custers J, van der Net J (1997): Functional Outcome in Osteogenesis Imperfecta : Disability Profiles using the PEDI . Pediatr Phys Ther 9: 18-22.

29. Haley S, Coster W, Ludlow L (1992): Pediatric Evaluation of Disability Inventory. New England Medical Center Hospitals.

30. Bleakney D, Kruse R (1996): Gross Motor Development and Children with Osteogenesis Imperfecta. Proceedings of Sixth International Conference on Osteogenesis Imperfecta, Netherlands.p.31.

31. Folio M, Fewell R (1983): Peabody Developmental Motor Scales and Activity Cards. Hingham, Mass. Teaching Resources Corp.

32. Cole B, Finch E, Gowland (1994): Physical Rehabilitation Outcome Measures. Toronto, Physical Therapy Association, Health and Welfare Canada Communications Group,Publishing, Supply and Services Canada.

33. Piper MC, Darrah J (1994): Motor Assessment of the Developing Infant. Philadelphia: WB Sanders.

34. Bayley N (1969): Manual for the Bayley Scales of Infant Development. New York: Psychological Corp.

35. Hughes JE (1979): Basic Gross Motor Assessment. Golden CO., Jeanne Hughes.

36. Bruinicks RH (1978): Bruinicks-Oseretsky test of motor proficiency: examiner's manual. United States of America: American Guidance Service Inc.

37. Chandler LS, Andrew MS, Swanson MW. (1980) Movement Assessment of Infants-a Manual. Rolling Bay, WA.98061: P.O. Box 4631.

38. Russell D, Rosenbaum P, Gowland (1993): Gross Motor Function Measure Manual. Hamilton, Ont. Gross Motor Measures Group, second edition.

39. Russell D, Rosenbuam P, Cadman D (1989): The Gross Motor Function Measure: a Means to Evaluate the Effects of Physical Therapy. Developmental Medicine and Child Neurology 31: 341-352.

40. Russell D, Palisano R, Walter S et al (1998): Evaluating motor function in children with Down syndrome: validity of the GMFM. Developmental Medicine and Child Neurology 40: 693-701.

41. Parker DF, Carriere L, Hebestreit H et al (1993): Muscle Performance and Gross Motor Function of Children with Spastic Cerebral Palsy. Developmental Medicine and Child Neurology 35: 17-23.

42. McLaughlan J, Bjornson K, Astley S et al (1991): Ability to detect functional change with the gross motor function measure: A pilot study. Developmental Medicine and Child Neurology 33 (Suppl.64), 26.

43. McLaughlin J, Bjornson K, Astley S et al ((1993): The role of selective dorsal rhizotomy in cerebral palsy remains to be established: Results of a pilot study. Manuscript to be submitted for publication.

44. Streiner D, Norman G (1996): Health Measurement Scales, A Practical Guide to their Development and Use. Oxford England. Oxford University Press, second edition, 108

45. Gross D, Conrad B (1994): Issues Related to the Reliability of Videotaped Observational Data. Western Journal of Nursing Research: 13:799-803.

46. Nunally JC (1978): Psychometric Theory. second edition, New York, McGraw-Hill.

47. Helmstadler GC (1964): Principles of Psychological Measurement. New York, Appleton-Century-Crofts.

48. Shrout PE , Fleiss JL (1979): Intraclass Correlation: Uses in assessing rater reliability. Psychological Bulletin 86: 420-428.

49. Fleiss JL (1986): The Design and Analysis of Clinical Experiments. John Wiley and Sons, New York.12 and 27.

50. Cohen JA (1960): Coefficient of agreement for nominal scales. Educ Psychol Meas 20:37.

51. Landis RJ, Koch GG (1974): The measurement of observer agreement for categorical data. Biometrics 33: 159-174.

52. Fleiss J, Cohen J (1973): The equivalence of weighted Kappa and the intraclass correlation coefficient as measure of reliability. Ed Psychol Meas. 33613-619.

53. Portnoy LG, Watkins MP (1993): Foundations of Clinical Research Applications to Practice. East Norfolk, Appleton and Lange.505-528.

54. Bjornson KF, Graubert CS, Buford, VL et al (1998): Validity of the Gross Motor Function Measure. Pedia Phys Ther 10:43-47.

55. Shapiro F (1985) Consequences of an Osteogenesis Diagnosis on Survival and Ambulation. J Ped Ortho 5: 456-462.

56. Nicholas R, James P (1990) Telescoping Intramedullary Stabilization of the Lower Extremities for Severe Osteogenesis Imperfecta. J Ped Ortho 10: 219-223.

57. Ryoppy S, Alberty A, Kaitila I (1987) Early Semiclosed Intramedullary Stabilization in Osteogenesis Imperfecta. J Ped Ortho 7:139-144.

List of Tables

Appendix 1

Gross Motor Measure	**Population**	**Format**	**Subscale**	**Standar-Dization**	**Purpose**	**Weakness For O.I. Population**
Alberta Infant Motor Scale (AIMS)	Infants at risk for motor problems, term to independent ambulation	Observation	Prone Supine Sitting Standing	Normal infants Term-18 months	Evaluate efficacy of rehab. Infants till independent ambulation	Valid only until onset of walking
Bayley Scale of Infant Development	Assess motor & mental devt. Birth-36 months	Task performance	Mental scale Psycho-Motor prone-stand,walk Jump, bal..	Normal infants 2-30 months	Clinical assessment and research	Not valid throughout adolescence
Peabody Developmental Motor Scale	0-83 months	Task performance	Gross motor reflexes,bal., non-locomotor Mobility & fine motor	Normal children	Detect change in children with disabilities	Valid until 83 months Some items unsafe for O.I.
Posture & Fine Motor Assessments of Infants (PFMAI)	2-6 months	Caregiver report	Posture & fine motor		Detect change in motor function	Valid until 6 months
Basis Gross Motor Assessment (BGMA)	Minor motor dysfunction 5.5-12.5 years	Task performance	Balance on one leg/ eyes closed, hop, skip, jump	Normal 6-12 years	Evaluate performance by quantifying quality of movt.	Valid only from 5.5-12.5. yrs. In mildly affected population
Movement Assessment of Infants (MAI)	0-12 months (adjusted) at high risk for motor dysfunction	Task performance	Muscle tone Reflexes Automatic reactions Volitional movt.		Screening tool Efficacy of physiotherapy.	Normal profile only for 4 & 8 months
Pediatric Evaluation of Disability Inventory (PEDI)	Children with motor disabilities ,func. Level less than 7 yrs.	Parent report Structured interview	Self care Mobility Social func. Care giver assistance	Normal children	Assess. Functional capacities and performance	Does not measure gross motor devt. Fit score problem
Bruinick-Oseretsky Test of Motor Profiency (BOTMP)	Mild motor dysfunction 4.5-14.5 yrs.	Task performance	Running speed Bal. Jumping jacks, pushups etc.	Normal children	Identify, evaluate motor dysfunction	Skills too risky Valid only from 4.5-14.5 yrs. Skills too difficult for Type III & IV
Gross Motor Function Measure (GMFM)	Cerebral palsy	Task performance	Lying& roll Sit,Crawl Stand, run		Evaluate efficacy of treatment	

GROSS MOTOR FUNCTION MEASURE GMFM

SCORE SHEET

Child's Name: _____ I.D. #: _____

Date of Birth: ____/____/____ (yy / mm / dd) Assessment date: ____/____/____ (yy / mm / dd)

Diagnosis: _____ Severity: ☐ ☐ ☐

Mild Moderate Severe

Evaluator's Name _____

Testing Conditions (e.g. room, clothing, time, others present)

_____

_____

_____

_____

The GMFM is a standardized observational instrument designed and validated to measure change in gross motor function over time in children with cerebral palsy.

*SCORING KEY		
	0	= does not initiate
	1	= initiates
	2	= partially completes
	3	= completes

*Unless otherwise specified, "initiates" is defined as completion of less than 10% of the item. "Partially completes" is defined as completion of 10% to less than 100%.

The scoring key is meant to be a general guideline. However, most of the items have specific descriptors for each score. It is imperative that the **guidelines be used for scoring each item.**

Contact address:
Dianne Russell, Gross Motor Measure Group, Chedoke-McMaster Hospitals, Chedoke Hospital, Building 74, Room 29, Box 2000, Station "A", Hamilton, Ontario L8N 3Z5

Children's Developmental Rehabilitation Programme at Chedoke-McMaster Hospitals, Hamilton, Ontario, Hugh MacMillan Rehabilitation Centre, Toronto, Ontario, and McMaster University, Hamilton, Ontario

(over)

Item	C: CRAWLING AND KNEELING	SCORE	
38.	PR: CREEPS FORWARD 6'	0 1 2 3	38.
39.	4 POINT: MAINTAINS, WEIGHT ON HANDS AND KNEES, 10 SECONDS	0 1 2 3	39.
40.	4 POINT: ATTAINS SIT ARMS FREE	0 1 2 3	40.
41.	PR: ATTAINS 4 POINT, WEIGHT ON HANDS AND KNEES	0 1 2 3	41.
42.	4 POINT: REACHES FORWARD WITH R ARM, HAND ABOVE SHOULDER LEVEL	0 1 2 3	42.
43.	4 POINT: REACHES FORWARD WITH L ARM, HAND ABOVE SHOULDER LEVEL	0 1 2 3	43.
44.	4 POINT: CRAWLS OR HITCHES FORWARD 6'	0 1 2 3	44.
45.	4 POINT: CRAWLS RECIPROCALLY FORWARD 6'	0 1 2 3	45.
46.	4 POINT: CRAWLS UP 4 STEPS ON HANDS AND KNEES/FEET	0 1 2 3	46.
47.	4 POINT: CRAWLS BACKWARDS DOWN 4 STEPS ON HANDS AND KNEES/FEET	0 1 2 3	47.
48.	SIT ON MAT: ATTAINS HIGH KN USING ARMS, MAINTAINS, ARMS FREE, 10 SECONDS	0 1 2 3	48.
49.	HIGH KN: ATTAINS HALF KN ON R KNEE USING ARMS, MAINTAINS, ARMS FREE, 10 SECONDS	0 1 2 3	49.
50.	HIGH KN: ATTAINS HALF KN ON L KNEE USING ARMS, MAINTAINS, ARMS FREE, 10 SECONDS	0 1 2 3	50.
51.	HIGH KN: KN WALKS FORWARD 10 STEPS, ARMS FREE	0 1 2 3	51.

TOTAL DIMENSION C ☐

Item	D: STANDING	SCORE	
52.	ON THE FLOOR: PULLS TO STD AT LARGE BENCH	0 1 2 3	52.
53.	STD: MAINTAINS, ARMS FREE, 3 SECONDS	0 1 2 3	53.
54.	STD: HOLDING ON TO LARGE BENCH WITH ONE HAND, LIFTS R FOOT, 3 SECONDS	0 1 2 3	54.
55.	STD: HOLDING ON TO LARGE BENCH WITH ONE HAND, LIFTS L FOOT, 3 SECONDS	0 1 2 3	55.
56.	STD: MAINTAINS, ARMS FREE, 20 SECONDS	0 1 2 3	56.
57.	STD: LIFTS L FOOT, ARMS FREE, 10 SECONDS	0 1 2 3	57.
58.	STD: LIFTS R FOOT, ARMS FREE, 10 SECONDS	0 1 2 3	58.
59.	SIT ON SMALL BENCH: ATTAINS STD WITHOUT USING ARMS	0 1 2 3	59.
60.	HIGH KN: ATTAINS STD THROUGH HALF KN ON R KNEE, WITHOUT USING ARMS	0 1 2 3	60.
61.	HIGH KN: ATTAINS STD THROUGH HALF KN ON L KNEE, WITHOUT USING ARMS	0 1 2 3	60:
62.	STD: LOWERS TO SIT ON FLOOR WITH CONTROL, ARMS FREE	0 1 2 3	62.
63.	STD: ATTAINS SQUAT, ARMS FREE	0 1 2 3	63.
64.	STD: PICKS UP OBJECT FROM FLOOR, ARMS FREE, RETURNS TO STAND	0 1 2 3	64.

TOTAL DIMENSION D ☐

TESTING WITH AIDS/ORTHOSES

ıdicate below with a check (✓) which aid/orthosis was used and what dimension it was first pplied. (There may be more than one).

ıid		Dimension	Orthosis		Dimension
ıollator/pusher	☐	_____	Hip Control	☐	_____
Valker	☐	_____	Knee Control	☐	_____
ı Frame crutches	☐	_____	Ankle-foot Control	☐	_____
;rutches	☐	_____	Foot Control	☐	_____
]uad Cane	☐	_____	Shoes	☐	_____
;ane	☐	_____	None	☐	_____
Vone	☐	_____	Other _____ (please specify)	☐	_____
Dther _____ (please specify)	☐	_____			

SUMMARY SCORE USING AIDS/ORTHOSES

DIMENSION	CALCULATION OF DIMENSION % SCORES	GOAL AREA (indicated with ✓ check)
A. Lying & Rolling	$\frac{\text{Total Dimension A}}{51} = \frac{\_\_\_\_}{51} \times 100 = \_\_\_\_\ \%$	A. ☐
B. Sitting	$\frac{\text{Total Dimension B}}{60} = \frac{\_\_\_\_}{60} \times 100 = \_\_\_\_\ \%$	B. ☐
C. Crawling & Kneeling	$\frac{\text{Total Dimension C}}{42} = \frac{\_\_\_\_}{42} \times 100 = \_\_\_\_\ \%$	C. ☐
D. Standing	$\frac{\text{Total Dimension D}}{39} = \frac{\_\_\_\_}{39} \times 100 = \_\_\_\_\ \%$	D. ☐
E. Walking, Running & Jumping	$\frac{\text{Total Dimension E}}{72} = \frac{\_\_\_\_}{72} \times 100 = \_\_\_\_\ \%$	E. ☐

TOTAL SCORE $= \frac{\%A + \%B + \%C + \%D + \%E}{\text{Total \# of Dimensions}}$

$= \frac{\_\_\_\_ + \_\_\_\_ + \_\_\_\_ + \_\_\_\_ + \_\_\_\_}{5} = \frac{\_\_\_\_}{5} = \_\_\_\_\ \%$

GOAL TOTAL SCORE $= \frac{\text{Sum of \% scores for each dimension identified as a goal area}}{\text{\# Goal areas}}$

$= \frac{\_\_\_\_}{\_\_\_\_} = \_\_\_\_\ \%$

CONSENTEMENT À LA PHOTOGRAPHIE MÉDICALE

Date_____

A l'égard des soins médicaux que reçoit_____, mon enfant ou l'enfant dont j'ai la garde, aux Hôpitaux Shriners pour l'enfant infirme, établissement hospitalier de_____, j'autorise par les présentes à ce que des photographies, diapositives et/ou bandes video, cinématographiques ou de télédiffusion soient prises de mon enfant ou de l'enfant dont j'ai la garde, ou de quelque(s) partie(s) de son corps, par les membres du personnel et le personnel qualifié de l'hôpital, pour les usages suivants et sous réserve des conditions suivantes:

(1) Les photographies pourront faire partie du dossier médical de mon enfant ou de l'enfant dont j'ai la garde et pourront être utilisées à titre d'exemple lors de cours ou conférences et faire partie de publications medicales, être publiées de quelque facon et en autant de fois que l'hôpital et les mémbres du personnel médical le jugeront opportuns;

(2) Toutes les mesures seront prises afin d'éviter l'identification du patient, si une partie de son visage ou une région de son corps généralement vêtue est en cause.

Je n'anticipe aucune compensation ou rémunération. La présente autorisation à quelqu'utilisation desdites photographies, diapositives et/ou bandes vidéo, cinématographiques ou de télédiffusion dégage expressément le photographe, le médecin traitant, l'hôpital et son personnel, les Hôpitaux Shriners pour l'enfant infirme et ses compagnies affiliées, Imperial Council, A.A.O.N.M.S., Shrine Temples, leurs officiers et membres, de toute responsabilité.

_____
Signature du parent ou tuteur

TEMOIN:

_____

Les Hopitaux Shriners **Etablissement de Montreal**

pour l'enfant infirme **Consentement à la photographie médicale.**

•Formulaire 1035A-1/84

CONSENT TO MEDICAL PHOTOGRAPHY

Date _____

In connection with the medical services which _____ , my child or ward, is receiving at the _____ Hospital Unit of the Shriners Hospitals for Crippled Children, I hereby consent that photographs, slides, television, videotape, or motion pictures may be taken of my child or ward or parts of his/her body by members of the staff and appropriate personnel of the hospital for the following uses and subject to the following conditions:

(1) That medical photographs may be a part of my child or ward's hospital record or used for illustrative purposes in lectures and medical publications, being published and republished in any manner which the hospital or medical staff shall deem proper;

(2) That every effort will be made to prevent personal identification, if any portion of the patient's face is the subject or if regions of the patient's body ordinarily covered by clothing are to be the subject of the photograph.

I expect no compensation or other remuneration. This consent as to any use of said photographs, slides, television, videotape or motion pictures shall act to expressly release from liability the photographer, the attending physician, the hospital and all its personnel, Shriners Hospitals for Crippled Children and affiliated corporations, Imperial Council, A.A.O.N.M.S., Shrine Temples, their officers and members.

_____

Signature of Parent or Legal Guardian

WITNESS:

_____

Shriners Hospitals for crippled children Unit

CONSENT TO MEDICAL PHOTOGRAPHY

• Form 1035

Appendix 4

Simple Kappa for items 3,4 and 19 of the Gross Motor Function Measure for Children with Osteogenesis Imperfecta

Item	raters	simple Kappa	95% confidence limits
3: Supine: lifts head 45 degrees	1x 2	0.730	0.447 -1.013
	1x4	0.638	0.333 -0.944
	1x5	0.552	0.250- 0.854
	2x4	0.913	0.752- 1.073
	2x5	0.824	0.611- 1.037
	4x5	0.826	0.620- 1.033
4: Supine: flexes right hip & knee through full range	1x3	0.890	0.682- 1.098
	1x4	0.890	0.682- 1.098
	1x5	0.890	0.682- 1.098
	3x4	0.774	0.478- 1.070
	3x5	1.000	1.000
	4x5	0.774	0.478-1.070
19: Supine: rolls to right side, attains sitting	1x4	0.712	0.419-1.005
	1x5	0.802	0.544-1.060
	4x5	0.712	0.407-1.017

Appendix 5

Weighted Kappa for items 3 and 19 of the Gross Motor Function Measure for Children with Osteogenesis Imperfecta

Item	raters	weighted Kappa	95% confidence limits
3. Supine: lifts head 45 degrees	1x2	0.796	0.447-1.013
	1x4	0.751	0.496-1.006
	1x5	0.656	0.396- 0.945
	2x4	0.949	0.852- 1.046
	2x5	0.850	0.649- 1.050
	4x5	0.903	0.782-1.023
19:Supine: rolls to right side, attains sitting	1x4	0.682	0.369- 0.995
	1x5	0.780	0.499- 1.061
	4x5	0.683	0.335- 1.010