Radiographic Measurements in Developmental Dysplasia of the Hip

Reliability and Validity of a Digitizing Program

Douglas R. Pedersen, PhD,*† Colleen A. Lamb, BS,* Lori A. Dolan, MA,† Heather M. Ralston, BS,* Stuart L. Weinstein, MD,† and Jose A. Morcuende, MD, PhD†

Abstract: A fundamental goal of research in developmental dysplasia of the hip (DDH) has been the identification of radiographic quantitative factors predictive of long-term outcome. The rapid development of digital imaging modalities opens the possibility for more systematic, automated measurement of radiographs. The authors have developed a software program for the study of DDH that integrates a user interface to select, store, and edit prescribed landmarks on digital radiographs. These points are used to construct and display reference lines analogous to those commonly drawn on physical films. Measurements are stored as a printable text file and as a data record. The results of this study demonstrated very good intrarater and interrater reliability coefficients and intermethod validity coefficients for the digitizing program. If this program is added to a software reporting system, measurements can be done on a real-time basis and can be included in the patient’s medical record.

Key Words: developmental dysplasia of the hip, reliability, validity, digitizing program, radiographic measurements

Materials and Methods

The radiographic files of 20 patients were randomly selected from 119 participants in a 30-year follow-up of DDH from the mid-1980s. An anteroposterior (AP) pelvic radiograph was selected and evaluated (Fig. 1) for each patient with regard to the acetabular angle of Sharp; the center–edge (CE) angle of Wiberg; the articulotrochanteric distance (ATD); the acetabular depth, width, and quotient; the spherical index of the femoral head and its width and height components; the percentage coverage of the femoral head; the femoral head-to-head ratio; and the maximum width of the teardrop.

For use with the digitizing program, the radiographs were scanned at 150 dots per inch (dpi) on an 8-bit (256 gray scale) whole-film scanner (Cobrascan 612T, Radiographic Digital Imaging Inc., Compton, CA) and archived to compact disc. The digitizing program, developed by one of the authors (D.R.P.) using PV-WAVE, prompts the user to mouse-click on 32 specific landmarks on each scanned image. The on-screen coordinates of these landmarks are then used to calculate the 12 radiographic measures using trigonometric and algebraic algorithms. All landmarks and algorithms were developed to correspond exactly to literature definitions of the radiographic measures. Both raters measured the 20 scanned radiographs in a single sitting for each occasion, using a Windows operating system with a 19-inch monitor at a resolution of 1,280 × 1,024 pixels, allowing for the full-screen display of each scanned image at a size reduction factor of two.

A fundamental goal of research in developmental dysplasia of the hip (DDH) has been the identification of radiographic quantitative factors predictive of long-term outcome, since a poor anatomic result is often detectable long before clinical symptoms appear. Ensuing studies of radiographic measurements have produced conflicting estimates of normal values and reports of low rater reliability. With such inconsistencies in research, the establishment of predictive factor(s) reliable enough to determine clinical course of treatment seems unlikely. To address these issues, we have developed a digitizing program to automate the measurement of 12 standard radiographic parameters, thereby reducing measurement time and allowing the automatic creation of a measurement database. The purpose of this study was to gather evidence for (1) the reliability of the digitizing program using intrarater and interrater reliability coefficients and (2) the validity of the digitizing program using intermethod validity coefficients comparing the digitizing program with traditional hand measurement.

Study conducted at the University of Iowa Hospitals and Clinics, Iowa City, Iowa.

From the *Department of Biomedical Engineering, College of Engineering, and the †Department of Orthopaedics and Rehabilitation, College of Medicine, University of Iowa, Iowa City, Iowa.

Financial support received from the Whitaker Foundation, the Children’s Miracle Network, and Mr. and Mrs. Herb and Nancy Townsend.

Reprints: Douglas R. Pedersen, Orthopaedic Biomechanics Laboratory, 2181 Westlawn, Iowa City, IA 52242 (e-mail: doug-pedersen@uiowa.edu).

Copyright © 2004 by Lippincott Williams & Wilkins

J Pediatr Orthop • Volume 24, Number 2, March/April 2004
The Sharp angle is measured from the pelvic horizontal to a line from the distal teardrop to the superolateral acetabular margin (the lateral rim) (Fig. 2A). To determine Wiberg’s CE angle, two lines are first constructed. The first line passes through the femoral head center and is perpendicular to the pelvic horizontal. The second line passes through the lateral rim and the femoral head center. The CE angle is the angle between these two lines (see Fig. 2B). The articulotrochanteric distance is the perpendicular separation between the superior trochanter and the superior femoral head, as measured along the femoral shaft midline (see Fig. 2C). The acetabular width is the distance from the inferior teardrop to the lateral rim (see Fig. 2D). The acetabular depth is the perpendicular distance from the midpoint of the width line to the acetabular dome (see Fig. 2D). The acetabular quotient is the ratio of acetabular depth to width, multiplied by 1,000. Femoral head coverage is the percent ratio between the horizontal distance from the lateral rim to the medial femoral head and the horizontal distance from the lateral femoral head to the medial femoral head (see Fig. 2E). The teardrop is the J-shaped boundary of the acetabular bed; the teardrop width is the horizontal separation between the teardrop margins (see Fig. 2F).

The evaluations were made by two raters, a staff pediatric orthopaedist (J.A.M.) and a medical student (C.A.L.), after agreeing on specific definitions for each radiographic measure. Rater J.A.M. measured the 20 radiographs on two occasions, once by hand and once using the digitizing program. Rater C.A.L. measured the 20 radiographs on three occasions, once by hand and twice using the digitizing program. For each measurement session, the order of evaluation was randomly varied with regard to patient, and the time between any two measurements for a given rater was 6 to 8 days.

Measurements by hand were made in two or three sittings on consecutive days. For standardization purposes, each rater used the same brand of sharpened soft lead pencil and combination protractor/ruler. Both raters also used identical data sheets to record measures for each patient. All markings on the AP radiographs were removed with alcohol pads between readings.

Reliability and validity were estimated with use of generalizability coefficients generated using the GENOVA software package. The mean square estimates used in the GENOVA program were computed with use of the GLM procedure in the SAS system. Only data for a single dysplastic hip in each patient were considered in the computation of these coefficients. Intrarater coefficients for the digitizer method and interrater coefficients for both the hand and digitizer methods were calculated to simulate a clinical situation in which one clinician evaluates a radiograph on a single occasion. The intermethod validity comparing hand and digitizing methods was assessed for each rater using intraclass correlation coefficients.
RESULTS

Numeric values for intrarater reliability, interrater reliability, and intermethod validity coefficients are shown in Table 1 (on a scale of 0 to 1, for none to exact match, respectively). All coefficients were greater than 0.85 with the exception of one or two values for acetabular width (WIDTH), femoral head sphericity width (FHSWIDTH), femoral head coverage (FMCOV), head-to-head ratio (HHR), and teardrop width (TDROP), and all or nearly all values for femoral head sphericity (FHS) and femoral head sphericity height (FHS-HEIGHT). Interrater reliability and intermethod validity coefficients were so low for FHS as to equal 0.000 in two instances.

In general, intrarater reliability coefficients (Fig. 3A) for the digitizer were greater than or equivalent to corresponding intrarater coefficients (see Fig. 3B) for the same method. Interrater reliability coefficients for the digitizer were in turn greater than or equivalent to corresponding intrarater coefficients for the hand method. The WIDTH measurement was an exception to these two trends, but the more widely reported acetabular quotient (AQUOT) measurement, calculated as a ratio of acetabular depth (DEPTH) to WIDTH, still followed the general patterns.

The intermethod validity coefficients showed a high correlation between hand and digitizer methods except for FHS and its two components (see Fig. 3C). Generally, validity coefficients for rater C.A.L. were greater than or equivalent to analogous coefficients for rater J.A.M.

DISCUSSION

Since the discovery of radiography more than 100 years ago, imaging studies have become the central component for the diagnosis and treatment in medicine in general, and especially in orthopaedic conditions. However, the exponential increase in the number and complexity of studies, together with limitations of space to store films and the management of the file rooms, have led to the destruction of study materials older than 7 years in many hospitals. These facts have had a tremendous impact on the study of pediatric orthopaedic conditions.
such as DDH, where the results of treatment in early childhood often cannot be seen until the child reaches maturity.

Advances in information systems and technology, in conjunction with economic forces driving improved reporting and integration of the imaging data into the patient medical record, are causing sweeping changes in the practice of radiology. Progress in digital photo technology and picture archiving and communication systems (PACS) have provided a means for dealing with those issues. A dramatic increase in computer power and network bandwidth and a decrease in the cost of archival storage have allowed large studies of many megabytes to arrive at display stations within seconds of examination completion. It is possible now to have secure, rapid access to all clinical information on patients, including imaging studies, anytime and anywhere. Therefore, we are witnessing a transformation of medical diagnostic imaging into the electronic environment, with a transition from a traditional film-based department to a near-filmless operation.

This rapid development of digital imaging modalities in medicine opens the possibility for more systematic, automated measurement in radiographic studies. However, software development in this area is still lagging behind. We are approaching this issue by developing a software program for the study of DDH. The program integrates a user interface to select, store, and edit prescribed landmarks on digital radiographs. These points are used by the program to construct and display reference lines analogous to those commonly drawn on physical films for protractor and length measurements. The femoral head center is located by overlaying an electronic Moss circle, under keyboard control. Measurements are stored as a print-
DDH. If this program is added to a software reporting system, measurements can be done on a real-time basis and can be included in the patient’s medical record. Future studies will take advantage of this technology for study of specific questions in this pathology.

REFERENCES


