Weighted Diagnostic Criteria for Developmental Dysplasia of the Hip

Andreas Roposch, MD, MSc, FRCS[1,2], Evangelia Protopapa, MSc[1], and Mario Cortina-Borja, PhD[1]

Objective To establish clinical diagnostic criteria for developmental dysplasia of the hip (DDH) that model the practices of expert clinicians.

Study design Of 23 clinical criteria for the diagnosis of DDH, ranked in order of diagnostic importance by international consensus, the 7 most highly ranked were placed in all possible combinations to create unique case vignettes. Twenty-six experts rated 52 vignettes for the presence of DDH. We modeled the data to determine which of the 7 criteria were associated with a clinician's opinion that the vignette represented DDH. From the resulting regression coefficients, for each vignette we calculated a probability of DDH. An independent panel rated the same vignettes using a visual analog scale response. We correlated the visual analog scale ratings with probabilities derived from the model.

Results Our model identified 4 of 7 criteria as predictive of DDH ($P < .001$): Ortolani/Barlow test ($\beta = 3.26$), limited abduction ($\beta = 1.48$), leg length discrepancy ($\beta = 0.74$), and first-degree family history of DDH ($\beta = 1.39$). There was substantial correlation between the probability of DDH predicted by the model and that derived from an independent expert panel ($r = 0.73$; $P < .001$).

Conclusion Weighted clinical criteria for inferring the likelihood of DDH produced consistent results in the judgment of 2 separate groups of experts. Using these weights, nonexperts could establish the probability of DDH in a manner approaching the practice of clinical experts. *(J Pediatr 2014; ■: ■-■).*

The diagnosis of developmental dysplasia of the hip (DDH) remains controversial, particularly during the first 8 weeks of life,[1,2] given the suggestion that many of the clinical findings considered indicative of DDH can resolve spontaneously.[3,4] Identification of findings that warrant follow-up or treatment is critical. Criteria obtained from clinical examination and patient history result in discharge, the need for repeat examination, or referral for ultrasound.[5] However, ultrasound is not a definite diagnostic test, especially during the first 8 weeks of life, because of its variable reliability,[6] the unresolved controversy regarding thresholds for defining DDH,[2] and the fact that many abnormalities resolve without treatment.[2,5] Ultrasonography can lead to overdiagnosis[5] by identifying abnormalities that will resolve spontaneously[7,8] or that will never cause symptoms.[5,9-12] Thus, clinicians need to make the most of the clinical context rather than order ultrasound for every patient.

Establishing conditional probabilities is critical for the diagnostic workup.[8] The likelihood of DDH in an individual with a positive Ortolani sign (near certainty) is not the same as that in an individual presenting with limited abduction (low).[8] Twenty-three different criteria for the diagnosis of DDH[13] have been reported, but without weighting, how these criteria should be used in daily practice remains unclear. The use of weighted diagnostic criteria has the potential to reduce the number of unnecessary ancillary ultrasound tests and reduce the risk of overdiagnosis or repeat but unnecessary imaging of hips with equivocal ultrasound findings. Other implications include a better understanding of true incidence of disease by reference to a uniform case definition; more accurate early diagnosis to avoid morbidities with late diagnosis; cost reductions by avoiding overreliance on ultrasound; and the ability to perform multicenter research with agreed-on definitions.

We have derived weighted diagnostic criteria that model the diagnostic practices of clinical experts. The use of these criteria should reduce the variability in assessments and management decisions for infants examined for the presence of DDH.

Methods

This study followed an item generation and consensus study, which we reported previously.[13] In brief, we previously surveyed all members of the European Pediatric Orthopaedic Society (EPOS) and other relevant societies, and performed a literature search and key informant interviews. EPOS is one of the largest professional societies of pediatric orthopedic surgeons, with more than 300 members from more than 30 countries. We elicited criteria considered important in the diagnosis of DDH in infants younger than 9 weeks, with DDH defined as a condition warranting treatment or follow-up with an orthopedic surgeon.

DDH	Developmental dysplasia of the hip
EPOS	European Pediatric Orthopaedic Society
VAS	Visual analog scale

From the [1]Institute of Child Health, University College London; and [2]Department of Orthopaedic Surgery, Great Ormond Street Hospital for Children, London, United Kingdom

Supported by the Bupa Foundation and Great Ormond Street Hospital Children's Charity. The authors declare no conflicts of interest.

http://dx.doi.org/10.1016/j.jpeds.2014.08.023

Table I. Item reduction process

Candidate items	Mean	Action	Final items
1. Ortolani test	8.7	Combine	1. Ortolani test, Barlow test
2. Barlow test	8.4	Combine	2. Abduction limitation
3. Asymmetry in abduction of ≥20°	7.5	Combine	Asymmetry in abduction of ≥20°
4. Any asymmetry of hip abduction	7.0	Combine	Abduction of both (or one) hips ≤45°
5. Breech presentation	6.7	Combine	3. Leg-length discrepancy (Galeazzi sign)
6. Abduction of both (or one) hips ≤45°	6.6	Combine	4. Torticollis
7. Leg-length discrepancy (Galeazzi sign)	6.5	Keep	5. Clubfoot or other fixed foot deformities
8. First degree relative was treated for DDH	6.1	Keep	6. Breech presentation or breech in last trimester
9. Breech positioning in-utero, cephalic presentation	5.5	Combine	7. First degree relative was treated for DDH
10. Family history of DDH, any relative	4.7	Drop	
11. Abduction of both (or one) hips ≤60°	4.5	Drop	
12. Oligohydramnios	4.2	Drop	
13. Female gender	4.1	Drop	
14. First born baby girl	3.9	Drop	
15. Abduction of both (or one) hips ≤70°	3.2	Drop	
16. Torticollis	3.1	Keep*	
17. Birth weight >4000 g	3.1	Drop	
18. Clubfoot or other fixed foot deformities	3.0	Keep*	
19. Asymmetry of groin or skin crease(s)	2.7	Drop	
20. Multiple birth or pregnancies of mother	2.7	Drop	
21. Born by cesarean section	2.6	Drop	
22. Postural (flexible) foot deformities	2.5	Drop	
23. Hip click in a stable hip	1.8	Drop	

Twenty-three candidate items were examined. Some were combined, kept as they were, or dropped, resulting in 7 final items.
*Items 16 and 18 were retained despite a low mean because historically they have been repeatedly associated with DDH.

A total of 188 items were pooled, and consensus was established using the Delphi technique, resulting in 23 clinical diagnostic criteria ranked in order of consensus-based mean value.[13] The criteria demonstrated clinical sensibility[14] and statistical homogeneity,[15] with a Cronbach α of 0.83.[13] The decisions about which of these 23 items to retain and possibly include in a final diagnostic index required striking a balance between comprehensiveness and feasibility.[14] We decided a priori to not consider items with a consensus-based mean value <5/10, with the exception of items 16 and 18; these were retained for further investigation because of their strong historical link to DDH. We also checked items for commonality and combined like items to create broader diagnostic constructs; for example, items 1 and 2 scored almost identically in the consensus and represent hip dislocation, and thus were combined; items 3, 4, and 6 concerned the diagnostic construct "restriction in hip abduction"; and items 5 and 9 related to the breech position (**Table I**).

Determining these criteria's relative importance in the diagnosis of DDH required examining the criteria when they occur together. Comprehensive combination of the 7 criteria resulted in $2^7 = 128$ computerized case vignettes. A case vignette is a structured description of a hypothetical patient consisting of sex, age, patient history, and clinical examination findings[16] (**Appendix**; available at www.jpeds.com).

We used hypothetical patients because actual patients would have introduced uncontrollable variables that would have biased the results. To validate the criteria, we recruited a group of clinical experts who had not participated in any of our previous studies. All of the surgeons participating in this study were professional members of EPOS or the Pediatric Orthopaedic Society of North America, board-certified, practicing in pediatric orthopedic institutions, treating DDH routinely, and English-speaking.

Participants were randomly allocated to binary or visual analog scale (VAS) response vignettes, including 26 surgeons using a binary response and an independent panel of 24 surgeons using a 10-cm VAS for each vignette. Each panel member evaluated 52 randomly chosen vignettes for the presence or absence of DDH. Using the binary responses, we developed a mixed-effects model, with a logit link and clusters defined by surgeon identity and a random effect on the intercept, with the 7 items as independent covariates. Fixed effects significant at the 5% level were retained in the final model. We used a quasi-likelihood information criterion[17] to compare goodness of fit. The quasi-likelihood information criterion confirmed that a model with vignettes nested within surgeons showed the best fit. From the resulting regression coefficients of the final model, we calculated a probability of DDH for each vignette. Our hypothesis was that if the coefficients were valid, then the probabilities of DDH for each vignette based on these regression coefficients would substantially correlate ($r > 0.70$)[18] with the judgments of the independent expert panel's VAS response. We used the Spearman correlation coefficient to describe the associations between the probabilities derived from the logistic regression model and the VAS response for all vignettes.

Results

Of the 23 standardized clinical diagnostic criteria, the item reduction process retained 4 items, combined 7 items, and eliminated 12 items, resulting in 7 final items for further evaluation (**Table I**). Of these 7 items, 4 contributed significantly