

RESEARCH ARTICLE

Inter-observer agreement between 2-dimensional CT versus 3-dimensional I-Space model in the Diagnosis of Occult Scaphoid Fractures

Tessa Drijkoningen, MD; Robert Knotter, MD; Emile G. Coerkamp, MD; Anton H.J. Koning, PhD;
Steven J. Rhemrev, MD, PhD; Frank J. Beeres, MD, PhD

Research performed at Surgery Department, Medisch Centrum Haaglanden, Den Haag, The Netherlands

Received: 12 October 2015

Accepted: 8 March 2016

Abstract

Background: The I-Space is a radiological imaging system in which Computed Tomography (CT)-scans can be evaluated as a three dimensional hologram. The aim of this study is to analyze the value of virtual reality (I-Space) in diagnosing acute occult scaphoid fractures.

Methods: A convenient cohort of 24 patients with a CT-scan from prior studies, without a scaphoid fracture on radiograph, yet high clinical suspicion of a fracture, were included in this study. CT-scans were evaluated in the I-Space by 7 observers of which 3 observers assessed the scans in the I-Space twice. The observers in this study assessed in the I-Space whether the patient had a scaphoid fracture. The kappa value was calculated for inter- and intra-observer agreement.

Results: The Kappa value varied from 0.11 to 0.33 for the first assessment. For the three observers who assessed the CT-scans twice; observer 1 improved from a kappa of 0.33 to 0.50 (95% CI 0.26-0.74, $P=0.01$), observer 2 from 0.17 to 0.78 (95% CI 0.36-1.0, $P<0.001$), and observer 3 from 0.11 to 0.24 (95% CI 0.0-0.77, $P=0.24$).

Conclusion: Following our findings the I-Space has a fast learning curve and has a potential place in the diagnostic modalities for suspected scaphoid fractures.

Keywords: 3D Imaging, Occult fracture, Scaphoid, Virtual Reality

Introduction

Hand trauma frequently occurs, especially in young male adults (1,2). When diagnosing scaphoid fractures, knowledge is needed of the anatomy and common fracture patterns of the bone. Specialized radiographic views or additional imaging is needed (3). To date no consensus regarding different diagnostic modalities has shown to be the gold standard (2).

An example of a diagnostic modality is the I-Space; a CAVE-like virtual reality system [Figure 1]. A "hologram" of a 3D modality dataset is visualized as a floating image in front of the viewers. The viewers wear a pair of glasses with polarizing lenses that allow the bony structures to be seen as free-floating objects in three dimensions. An advantage is that a virtual pointer can be used to interact with this "hologram". This makes it possible to 'slice' through the hologram and to perform all kinds of measurements. In other fields very promising results are shown using this method (4-10).

In this pilot study CT-scans of patients suspected of an occult scaphoid fracture, were analyzed using the I-Space. The aim of this study is to analyze the inter-observer agreement of the current most used modality (CT-scan (2D view)) compared to virtual reality (3D view) in diagnosing occult scaphoid fractures.

Materials and Methods**Design**

This study used data from a prior study, a convenient sample comprised of 24 of 149 patients, with a high clinical suspicion of a scaphoid fracture but no fracture on radiograph. All patients had undergone a CT-scan, at the Emergency Department of one hospital. In all patients the CT scan was performed within 24 hours after presentation according to a standardized protocol (1). The 2D CT-scans were initially assessed for fractures of the scaphoid by 2 physicians and 2 residents. CT-scans with slice thickness >1.5 mm will not produce a reliable

Corresponding Author: Tessa Drijkoningen, Massachusetts General Hospital, Orthopaedic Hand and Upper Extremity Service, Yawkey Center, Suite 2100, 55 Fruit Street, Boston, MA 02114, USA

Email: tessa.drijkoningen@gmail.com



THE ONLINE VERSION OF THIS ARTICLE
[ABJS.MUMS.AC.IR](http://abjs.mums.ac.ir)

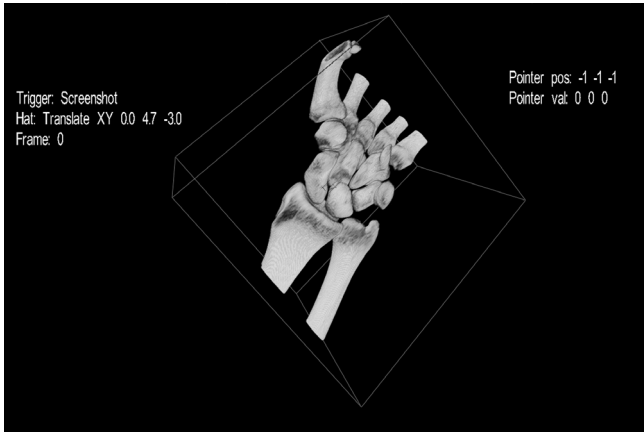


Figure 1. Impression of the I-space concept, a virtual reality system. Using projectors and polarizing glasses to get a three-dimensional interactive hologram created in front of the observer.

3D reproduction in the I-Space, therefore hundred twenty patients were excluded because CT-scan images had a slice thickness of more than 1.5 mm. Five patients were excluded as they had multiple fractures in the carpal bones. In this study the images were assessed by seven observers in the I-space (a CAVE-like virtual reality system) after conversion.

To measure inter-observer reliability seven observers viewed the blinded CT scans in the I-Space and were asked to diagnose the presence of a scaphoid fracture [Figure 2]. Three of the seven observers also assessed the 2D CT-scans 5 years before this study. Three of the seven observers assessed the CT-scans twice in the I-Space. Scans were randomized the second time and time between the sessions was at least 3 weeks.

Patients gave their written consent and permission in the original studies to use their data for future studies and institutional approval for the use of study data was granted by the Medical Review Board of our institute.

Radiology

The Computed radiography system; General Electric Lightspeed Qx/GT 64-slice Scanner Pewaukee, WI was used. The patient was lying on the scanner couch with the involved hand extended forward palm down. The wrist was laying in neutral flexion and neutral radial ulnar deviation (1). CT was performed by standard protocol (1). Some patients had to be scanned with their cast on for clinical reasons. The CT-images were transferred to the I-Space (Barco NV, Kortrijk, Belgium), at the department of Bioinformatics, Erasmus MC (Rotterdam, Netherlands).

The I-Space is a four-walled CAVE TM-like virtual reality system. Eight projectors create 3D images that can be seen using a set of polarizing glasses. The V-Scope volume rendering application creates an interactive "hologram" of the CT-dataset and is used to perform measurements. The dataset can be manipulated (e.g., rotated, resized) interactively using a wireless joystick projecting a virtual pointer. The user can turn on a clipping plane that is attached to the virtual pointer, thereby allowing intuitive placement. In addition the dataset can be rotated while

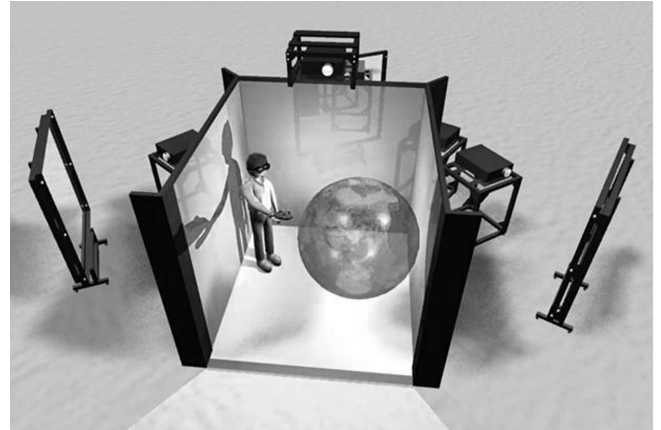


Figure 2. Visualization of the scaphoid in three dimensions as viewed in the I-space.

manipulating the clipping plane, making it very simple to investigate suspected areas from different angles. The depth perception offered by the I-Space enhances the visualization of fractures. Finally the transparency and contrast of bone as well as soft tissue can interactively be changed, which may increase the visibility of fractures.

Statistical analysis

A consensus of 2 physicians and 2 training physicians for the 2D assessment of the CT-scan was used as the gold standard in this study (1). The observer demographics and patient characteristics were described with frequencies and percentages for categorical variables and with mean \pm standard deviation and range for continuous variables. The kappa value was calculated for inter- and intra-observer agreement, also it was determined whether a second assessment in the I-Space correlated with improvement in intra-observer agreement. Kappa is a quantitative measure of agreement among observers and takes into account that observers will sometimes choose the same answer to a question by chance (10). A perfect agreement among observers would be reflected as a kappa of 1, whereas agreement totally based on chance would equate to a kappa of 0. Interpretation of kappa often is done by a classification by Landis and Koch in which a kappa between 0.01 and 0.20 is considered to reflect slight agreement, a value between 0.21 and 0.40 as fair agreement, between 0.41 to 0.60 as moderate agreement, 0.61 and 0.80 as substantial agreement, and greater than 0.81 as almost perfect agreement (12).

Baseline characteristics

Twenty-four patients were included from two prior studies. No Patients had signs of a fracture on conventional X-ray. A total of 3 (12.5%) patients were diagnosed with definite fractures of the scaphoid on the CT-scan. The mean age for all patients was 36 years (range 18-60) with 14 (58.3%) males and 10 (41.7%) females [Table 1].

Results

A total of seven observers assessed the CT-images of all patients in the I-Space. Two of the observers were a

Table 1. Observer Demographics

	obs 1	obs 2	obs 3	obs 4	obs 5	obs 6	obs 7
Male/ Female	male	male	female	male	male	male	male
Certified Physician/ Resident	physician	resident	resident	resident	resident	physician	resident
Years in practice	to 15 10	to 10 5	to 5 0	to 5 0	to 10 5	to 25 20	to 5 0
Specialty	surgery	surgery	radiology	surgery	surgery	radiology	surgery

Table 2. Patient characteristics (n=24)

Characteristic	Number
Mean age (range)	36 (18-60)
Male sex (%)	14 (58.3 %)
Positive scinigraphy (%)	6 (25 %)
Positive CT-scan (%)	3 (12.5 %)

radiologist and a resident radiology, one was a surgeon and four were surgery residents [Table 2]. There was a variable range of years in practice. The Kappa value for the agreement for each observer assessing the images in the I-space versus the result of the CT-scan, conducted from assessment of 4 observers in the previous study was calculated. The lowest Kappa value measured was 0.11, the highest value was 0.33 for the first assessment [Table 3].

The three observers that reviewed the images for a second time in the I-Space showed greater level of agreement between 2D CT and 3D I-Space models: the kappa value for observer 1 improved from 0.33 to 0.50 (95% CI 0.26-0.74, $P=0.01$), observer 2 from 0.17 to 0.78 (95% CI 0.36-1.0, $P<0.001$), and observer 3 from 0.11 to 0.24 (95% CI 0.0-0.77, $P=0.24$) [Table 4].

The Fleiss Kappa for inter-observer agreement for all observers in the I-space was 0.20 (95% CI 0.07-0.39, $P=0.00$). When analyzing the subset of observers that performed a second assessment in the I-space it is shown that inter-observer agreement improved from 0.01 (95% CI 0.0-0.36, $P=0.252$) for the first assessment to 0.37 (95% CI 0.12-0.63, $P=0.001$) for the second assessment [Table 5].

Discussion

The I-Space is a fairly new 3D-display modality, which is

Table 3. Level of agreement I-space vs CT-scan

	Kappa (1st assessment) (95% CI, p value)	Kappa 2nd assessment) (95% CI, p value)
Observer 1	0,33 (0,0-0,78, p=0,08)	0,50 (0,26-0,74, p=0,01)
Observer 2	0,17 (0,0-0,65, p=0,41)	0,78 (0,36-1,0, p<0,001)
Observer 3	0,11 (0,0-0,55, p=0,57)	0,24 (0,0-0,77, p=0,24)
Observer 4	0,11 (0,0-0,40, p=0,44)	n.a.
Observer 5	0,17 (0,0-0,65, p=0,41)	n.a.
Observer 6	0,33 (0,0-0,77, p=0,08)	n.a.
Observer 7	0,33 (0,0-0,91, p=0,09)	n.a.

rarely used in clinical settings. This pilot study using this 3D-display modality was performed, as there is no golden standard in diagnosing scaphoid fractures (2). Results of this study suggest that the I-space is a modality with a fast learning curve and a potential clinical usefulness to assess occult fractures.

Previous studies point out that CT-scans are comparable in diagnosing scaphoid fractures to MRI and that in both modalities fractures of the scaphoid can be missed (2). Imaging modalities such as the I-Space were until now mostly used in research setting and for soft tissue studies (4-7,9,10). One prior study was published using the I-Space focusing on postoperative anterior cruciate ligament reconstruction. They found the best inter-observer agreement when using the I-Space compared to normal CT-views (3).

The first limitation is the fact that there is no gold standard in diagnosing scaphoid fractures. In this study we compared the results in the I-Space with the results from the CT-scans. The CT-scans were performed for use in a prior study, consequently 125 patients had to be excluded due to incomplete saving of all CT slices or the saved slices were too thick. This results in loss of power, however for this pilot study the power remains adequate. No systematic bias was introduced by excluding these patients, for quality of the scans was the selection criterium. The quality of some of the images was low as a number of scans were performed while the patients arm was in a plaster cast. As this was the same for the 2D viewed CT-scans we assume that this should not make a difference. Making CT-scans without a cast is recommended in general, as less radiation is needed.

How can the present results be explained? None of the observers had previously assessed images in the I-Space. As a result, the level of agreement for the first assessment in the I-Space with conventional CT-images is poor (kappa range 0.11 – 0.33), the intra-observer level of agreement for the I-space was poor as well (Fleiss kappa 0.20). All patients included in this study had a clinical suspicion of an occult fracture not visible on radiographs, making images difficult to assess in any modality. Furthermore, previous literature shows variable intra- and inter-

Table 4. Level of agreement second assessments I-space

	Kappa (95% CI, p value)
Observer 1	0,25 (0,0-0,69, p=0,21)
Observer 2	0,25 (0,0-0,76, p=0,19)
Observer 3	0,41 (0,0-0,87, p=0,04)

observer reliability for CT-scan as the primary imaging modality in optimal conditions as well (13). Amongst the 3 observers who viewed the images twice in the I-Space the intra-observer agreement rose significantly for the second assessment. This trend was also shown in a comparison of the inter-observer reliability in these 3 observers (0.01 to 0.37). These results suggest that the I-space is an easy modality to learn which could be good for teaching purposes. At this point it is difficult to predict the exact diagnostic value as this is a fairly new modality for fracture diagnosis.

What are the clinical implications? Viewing CT-images in the I-Space 3D-display might be useful in pre-operative planning and teaching settings as the images help picturing a complete 3D view. Anatomical details can be realistically pictured on a great scale in the I-Space, resulting in a very illustrative overall picture for training physicians. For the I-Space method a separate room with projectors is needed, consequently it might not be feasible in every hospital setting. Although in teaching hospitals it could be of worthwhile investment. The system is user friendly and will not take much time getting used to. These advantages might make this a new progressive imaging device in clinical settings.

The aim of this study was to get pilot data on the diagnostic value of the I-Space in detecting occult scaphoid fractures. Following our findings the I-Space has a fast learning curve. We carefully recommend the use of the I-Space in pre-operative planning and for teaching purposes. Our recommendation is do more studies

using the I-Space with a separate scanning-protocol and optimal image quality when searching for the diagnostic value in occult scaphoid fractures. Moreover would it be worthwhile to study the value in pre-operative planning and teaching for physicians.

Tessa Drijkoningen MD
Orthopaedic Hand and Upper Extremity Service,
Massachusetts General Hospital, Boston, USA

Robert Knotter MD
Emergency Department, Reinier de Graaf Gasthuis, Delft,
The Netherlands

Emile G. Coerkamp MD
Radiology Department, Medisch Centrum Haaglanden,
Den Haag, The Netherlands

Anton H.J. Koning PhD
Department of Bioinformatics, Erasmus MC University
Medical Center, Rotterdam, Netherlands

Steven J. Rhemrev MD PhD
Frank J. Beeres MD PhD
Surgery Department, Medisch Centrum Haaglanden, Den
Haag, The Netherlands

References

1. Rhemrev SJ, de Zwart AD, Kingma LM, Meylaerts SA, Arndt JW, Schipper IB, et al. Early computed tomography compared with bone scintigraphy in suspected scaphoid fractures. *Clin Nucl Med.* 2010; 35(12):931-4.
2. Mallee WH, Wang J, Poolman RW, Kloen P, Maas M, de Vet HC, et al. Computed tomography versus magnetic resonance imaging versus bone scintigraphy for clinically suspected scaphoid fractures in patients with negative plain radiographs. *Cochrane Database Syst Rev.* 2015; 5(6):CD010023.
3. Hannemann PF, Brouwers L, Dullaert K, van der Linden ES, Poeze M, Brink PR. Determining scaphoid waist fracture union by radiographic examination: an analysis of reliability and validity. *Arch Orthop Trauma Surg.* 2015; 135(2):291-6.
4. Timmermans L, Deerenberg EB, van Dijk SM, Lamme B, Koning AH, Kleinrensink GJ, et al. Abdominal rectus muscle atrophy and midline shift after colostomy creation. *Surgery.* 2014; 155(4):696-701.
5. Meuffels DE, Potters JW, Koning AH, Brown CH Jr, Verhaar JA, Reijman M. Visualization of postoperative anterior cruciate ligament reconstruction bone tunnels: reliability of standard radiographs, CT scans, and 3D virtual reality images. *Acta Orthop.* 2011; 82(6):699-703.
6. van Adrichem LN, van Vlimmeren LA, Cadanová D, Helders PJ, Engelbert RH, van Neck HJ, et al. Validation of a simple method for measuring cranial deformities (plagiocephalometry). *J Craniofac Surg.* 2008; 19(1):15-21.
7. Baken L, van Gruting IM, Steegers EA, van der Spek PJ, Exalto N, Koning AH. Design and validation of a 3D virtual reality desktop system for sonographic length and volume measurements in early pregnancy evaluation. *J Clin Ultrasound.* 2015; 43(3):164-70.
8. Baken L, Rousian M, Koning AH, Bonsel GJ, Eggink AJ, Cornette JM, et al. First-trimester detection of surface abnormalities: a comparison of 2- and 3-dimensional ultrasound and 3-dimensional virtual reality ultrasound. *Reprod Sci.* 2014; 21(8):993-9.
9. Rifouna MS, Reus AD, Koning AH, van der Spek PJ, Exalto N, Steegers EA, et al. First trimester trophoblast and placental bed vascular volume

- measurements in IVF or IVF/ICSI pregnancies. *Hum Reprod.* 2014; 29(12):2644-9.
10. Baken L, van Heesch PN, Wildschut HI, Koning AH, van der Spek PJ, Steegers EA, et al. First-trimester crown-rump length and embryonic volume of aneuploid fetuses measured in virtual reality. *Ultrasound Obstet Gynecol.* 2013; 41(5):521-5.
 11. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med.* 2005; 37(5):360-3.
 12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977; 33(1):159-74.
 13. Adey L, Souer JS, Lozano-Calderon S, Palmer W, Lee S, Ring D. Computed tomography of suspected scaphoid fractures. *J Hand Surg Am.* 2007; 32(1):61-6.