Potential sources of quantification error when retrospectively assessing metacarpal bone loss from historical radiographs by using digital X-ray radiogrammetry: an experimental study

Johan Kälvesten, Torkel B. Brismar and Anders Persson

Linköping University Post Print

N.B.: When citing this work, cite the original article.

Original Publication:
http://dx.doi.org/10.1016/j.jocd.2013.04.002

Copyright: Elsevier
http://www.elsevier.com/

Postprint available at: Linköping University Electronic Press
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-105587
Potential sources of quantification error when retrospectively assessing metacarpal bone loss from historical radiographs by using digital X-ray radiogrammetry, an experimental study.

J Kälvesten\textsuperscript{1,2,3}, T B Brismar\textsuperscript{4}, A Persson\textsuperscript{1,2}

\textsuperscript{1} Radiology, Department of Medicine and Health Sciences, Faculty of Health Sciences, Linköping University, Östergötland County Council, University hospital, Linköping, Sweden

\textsuperscript{2}Center for Medical Image Science and Visualization (CMIV), Linköping University, Sweden

\textsuperscript{3}Sectra Imtec AB, Linköping, Sweden

\textsuperscript{4}Karolinska Institutet, Department for Clinical Science, Intervention and Technology, Division of Radiology, Karolinska University Hospital, Sweden

Correspondence:

Johan Kälvesten, Källgatan 24, 586 62 Linköping, Sweden.

Email: johan.kalvesten@gmail.com
Abstract

During the past 15 years, digital X-ray radiogrammetry (DXR) has been used to measure metacarpal bone mineral density (BMD). BMD is often measured in existing cohorts where X-ray images were not acquired in accordance with the DXR imaging protocol (DIP). The purpose of the current study was to analyze how deviations from DIP in historical radiographs may affect the reproducibility of DXR-BMD measurements. Cadaver hand phantoms were used to conduct repeat measurements of deviations from DIP with respect to voltage, exposure, lateral displacement, supination, combination of lateral displacement and supination or rotation, extension of the wrist, and edge enhancement. Direct digital radiography (Aristos; Siemens Healthcare, Erlangen, Germany) was used for image acquisition, and dxr-online (Sectra, Linköping, Sweden) for DXR-BMD measurements. The influence of the tested deviations from DIP ranged from 0 mg/cm² to 32.5 mg/cm² (0–6.8%). On repetition with the same specimen, none of the deviations resulted in a within-specimen reproducibility error greater than 2 mg/cm² (0.4%, equivalent to a T-score of 0.042). Among the tested deviations, all except tube voltage had a magnitude greater than the normal measurement noise for the technique and must therefore be considered when planning a study based on historical images.

Keywords

Bone mineral density; digital X-ray radiogrammetry; rheumatoid arthritis; osteoporosis
Introduction

Evaluation of bone damage, such as erosions, in radiographs of hands and feet is often a standard procedure when diagnosing rheumatoid arthritis (RA) and evaluating disease progress and treatment effect (1). The development of digital X-ray radiogrammetry (DXR) has made it possible to measure bone mineral density (BMD) by analyzing hand radiographs. The technique was often not available, or taken into consideration, when the radiographs were obtained. During the past decade, rheumatologists have become increasingly interested in retrospectively evaluating historical radiographs to study the rate of BMD loss related to different treatment regimens (2,3,4,5,6,7). This could provide valuable additional information from well-controlled historical clinical trials and enable many clinical studies that would otherwise have been unfeasible.

DXR-BMD measurements are highly reproducible when conducted in accordance with device labeling. Their coefficient of variation (CV%) of about 0.4% (8,9) compares favorably with the approximately 0.9% of dual energy X-ray absorptiometry (DXA) of the hand (10), the 0.7% of computed tomography (CT) (11) and the 1-3% of DXA hip (12). The high reproducibility of DXR makes it suitable for evaluating treatment effect, because very small changes in BMD can be detected with a high statistical power.

For clinical use, adherence to a DXR imaging protocol (DIP) with specific imaging parameters is required. However, radiographs for evaluating RA in historical clinical trials or for normal clinical use have often not been obtained according to DIP. It is also likely that image acquisition parameters may vary between, and within, centers when a study spans several years, rendering the reproducibility of BMD measurements from historical radiographs sub-optimal. However, advantages in study time, feasible follow-up time, study size and cost may compensate for those drawbacks, but the limitations must be properly understood when analyzing the data obtained. Our previous analysis of historical cohorts has shown that the following deviations from DIP are common. tube voltage, exposure, both hands in a
single exposure (i.e. hand displacement), supination of the hand, extension of the wrist, and image post processing. The influence of certain common imaging parameters in historical cohorts has been studied previously (13), but more knowledge is required on the impact of positioning.

The purpose of the current study was to analyze how a variation in imaging parameters and hand positioning in historical radiographs may affect the reproducibility of DXR-BMD measurements.

Material and methods

Digital X-ray Radiogrammetry

Digital X-ray radiogrammetry is a digital version of the traditional technique of radiogrammetry (14). The technique has been described in detail previously (8, 15). Briefly, this method provides a BMD estimate in g/cm² based on an automated analysis of the geometry and texture of cortical bone in the hand. Three measurement regions are automatically placed around the narrowest point of metacarpals 2, 3 and 4 (Figure 1). The periosteal and endosteal edges of the cortical bone are delineated within the measurement regions, and the cortical bone volume per area (VPA) is estimated by modeling the cortical bone as the difference between two concentric cylinders. DXR-BMD is then defined as DXR-BMD=\(c \times VPA_{\text{comb}} \times (1-P)\), where P is the estimated 3D porosity of the cortical bone, and c is an empirical density constant defined through a calibration of the specific acquisition modality type.

Study material

Anthropomorphic cadaver hand phantoms (The Phantom Laboratory, Salem, NY, USA) from five different individuals were used (Figure 2). The phantoms are constructed from a natural human skeleton cast
inside a material designed to have attenuation equivalent to soft tissue. The five phantom specimens ranged in BMD from 0.35 to 0.60 g/cm², corresponding to T-score values of -4.9 to 0.5.

Only one of the phantoms, with a BMD of 0.60 g/cm², had a physical shape allowing the wrist to extend 8° (variation 7, below).

**Imaging parameters**

The radiographs were obtained using an Aristos direct digital X-ray modality with a Trixell 4600 detector (Philips Medical Systems, Eindhoven, The Netherlands). The imaging of the five cadaver hands was repeated using ten different settings of the imaging parameters as listed below, resulting in a total of 100 radiographs. DXR-BMD was analyzed using Sectra dxr-online (Sectra, Linköping, Sweden). The imaging parameters obtained were:

1. according to the standard protocol (50 kV, 3.2 mAs) (3 repeats)
2. increased voltage (from 50 kV to 52 kV) (2 repeats)
3. decreased exposure (from 3.2 to 2.0 mAs) (2 repeats)
4. lateral displacement of 6.6 cm, simulating two hands on the same radiograph (2 repeats)
5. elevation of the thumb (i.e. supination of the hand) 1 cm, simulating joint inflammation preventing the palm from being positioned flat on the detector (2 repeats)
6. elevation of the thumb 2 cm (Fig 2) (2 repeats)
7. extension of the wrist 8°, (1 phantom, 5 repeats)
8. lateral displacement of 6.6 cm combined with 6° rotation (2 repeats)
9. lateral displacement of 6.6 cm combined with thumb elevation 1 cm (2 repeats)
10. standard hand examination settings used at the local medical center, including the standard edge enhancement processing used at this center (2 repeats)
Statistics

The impact of the different imaging variations on DXR-BMD was evaluated by calculating the variation in mean absolute DXR-BMD normalized to the average DXR-BMD for each phantom specimen under DXR protocol conditions using Equation 1, below

\[
\frac{\sum_{ph=1}^{N_{ph}} \sum_{i=1}^{N_{rep}} |BMD_{ph,i} - \mu_{ph}|}{N_{ph} \cdot N_{rep}}
\]

Eq. 1, where

- \( N_{ph} \): Number of phantoms.
- \( N_{rep} \): Number of repetitions per phantom.
- \( BMD_{ph,i} \): Measured DXR-BMD of phantom repetition \( i \) under imaging variation.
- \( \mu_{ph} \): Average DXR-BMD for phantom under DXR protocol conditions.

F-test of equality of variances and Welch’s two-sample t-test were used to evaluate significance. The impact of specific acquisition conditions on reproducibility was evaluated by within-phantom standard deviation. The hypothesis that acquisition conditions affected all phantom specimens to the same degree was tested with ANOVA F-tests. P values < 0.05 were considered significant.

Results

When using the standard imaging protocol, the average BMD of the five hands was 0.475 g/cm² (SD=0.10, range 0.349 to 0.604 g/cm²). All evaluated changes in image acquisition conditions, except voltage, affected the DXR-BMD significantly (Table 1, Figure 3). The 2 cm supination had the greatest influence on DXR-BMD, with a mean absolute deviation of 9.2 mg/cm² (SD=13.0, range -32.5 to 4.1 mg/cm²). Application of edge enhancement had a similar impact on the DXR-BMD measurement to that of the 2 cm supination, on average 9.1 mg/cm² (SD=4.2 range -16.1 to -3.1 mg/cm²). An increase in tube
voltage from 50 kV to 52 kV did not significantly affect the DXR-BMD (average 0.6 mg/cm², range -2.1 to 1.3 mg/cm²).

When the capture deviation was repeated with the same specimen, the maximum within-specimen reproducibility error was smaller than 2 mg/cm² (Table 2). In one case (a repeat), DXR measurement was not possible after 2 cm supination because the metacarpals were superimposed.

For all image capture protocol deviations, except increased voltage and (borderline) reduced exposure, the magnitude of the BMD measurement error was significantly dependent on the cadaver specimen (data not shown).

**Discussion**

When historical radiographs are used, several parameters may affect DXR-BMD. As long as the parameters are constant during a study, they do not affect the evaluation of changes in BMD (Table 2). However, when the parameters are changed, or varied, during a study that may span a decade, they are likely to have an impact on measurement reproducibility that must be taken into account. It is seldom possible to identify variation in exposure and voltage retrospectively. However, in this study we observed only a limited effect of exposure, an average of 3.4 mg/cm² (0.72%) for a 37.5% reduction of mAs (2 mAs), and no statistically significant impact of voltage offset of 2 kV. This corresponds well with findings by Böttcher et al, who observed an SD of 3 mg/cm² for an exposure range of 3–8 mAs and an SD of 8 mg/cm² when a wide voltage range of 40–52 kV was used when analyzing analog radiographs (13).

Variations in positioning had a greater influence on the DXR-BMD, but they are easier to identify retrospectively. The positioning error that had the greatest influence on DXR-BMD was severe supination of the hand. Such an elevation can be observed when the patient does not extend the thumb or flatten the hand properly. Based on our experimental study, a change in supination between serial radiographs
corresponding to a 2 cm elevation of the thumb would result in an average absolute effect of 9.2 mg/cm² (1.9%, corresponding to a T-score of 0.19). This can be compared with DXR-BMD change rates of +2 to -10 mg/cm²/month that can be observed in individual RA patients (8). Approximately 10–20% of the RA patients have a DXR-BMD loss rate of -2.5 mg/cm²/month or more (6,8), depending on patient demographics, treatment etc.

When the purpose of a retrospective study is to assess the risk of osteoporosis, and not the changes in DXR-BMD over time, the impact from deviation in image acquisition from the DXR device reference protocol is the limiting factor. A 1 mg/cm² offset corresponds to 0.021 units in T-score for Caucasian women (16). The observed impact of all the tested deviations from the DXR imaging protocol on T-score, ranging from 0 to 0.2, is moderate in relation to the normal variation in DXR-BMD T-score in the healthy population, even when adjusted for age, sex and ethnicity.

However, when planning a retrospective study to evaluate longitudinal change in DXR-BMD based on the analysis of non-DXR-protocol images, the level of supination and other positioning changes between serial images in the cohort should be evaluated to estimate their potential impact. For example, the feasibility of a retrospective study of treatment effect on DXR-BMD could be assessed by first estimating the expected treatment effect on DXR-BMD and then relating that to the known capture deviations between baseline and follow-up images in the cohort. Take a study with 50 subjects per study arm and 3 months between the radiographs. If the treatment effect is estimated as 2.1 mg/cm² (equivalent to what was observed in the BeSt trial (4)), the maximum allowable reproducibility error of DXR-BMD, including natural variation and measurement error, would be 4.2 mg/cm² (STATA 12.1, two-sample, one-sided) if a desired alpha error level of 0.05 and statistical power of 80% is applied. Tables 1 and 2 show that whereas such a study would be feasible if there is consistent supination, wrist extension and edge
enhancement between baseline and follow-up images, it would not have the desired power if most of the cases significantly changed any of those factors.

There was a difference between the individual cadaver specimens in how much the DXR-BMD measurement was affected by changes in positioning. We did not observe any substantial association between the degree to which a phantom was affected by a positioning change and the absolute level of DXR-BMD of the phantom. We surmise that the difference in effect between phantoms was due to differences in the anatomy between different specimens, e.g. variations in shape and striations in the metacarpal bones.

A limitation of our study was that the bones in the cadaver phantoms were fixed in their positions. Thus, the metacarpal bones probably lacked the rotation and movement patterns they would have had in vivo. Another limitation is the number of samples: 5 cadaver specimens for most deviations and only 1 for wrist extension.

In one case (a repeat), DXR measurement was not possible after 2 cm supination. This was caused by the combination of imperfect positioning of the metacarpals within the cadaver phantom and the superpositioning of the metacarpals.

One of the image series in this study included the standard edge enhancement used for hand radiographs at the local medical center. Considering that the type and implementation of the edge enhancement has a variable influence on the edge tracing used in DXR, this series cannot be considered representative of all types of edge enhancement. It merely indicates how a specific common type of edge enhancement can influence DXR-BMD measurements.

The current study is valid only for off-label retrospective use of DXR-BMD and does not consider its prospective on-label use.
This study provides information about how off-label variations in positioning and imaging parameters affect the reproducibility of DXR analysis of historical radiographs. We conclude that all the tested imaging protocol deviations, except tube voltage, for one or more of the five specimens affected the DXR-BMD measurement to a degree that was greater than the normal measurement reproducibility of DXR-BMD. When designing and planning a study based on non-DXR-protocol images, an analysis should be made to evaluate which image acquisition variations are acceptable and which are not. The estimated variations in DXR-BMD obtained in this experimental study may be used for such calculations.
Acknowledgements

We thank Monica Nyberg at the Department of Radiology, Linköping University Hospital, for assistance in acquiring the X-ray images. Sectra has provided the DXR measurements at no charge.

References

Figure Legends

**Figure 1:** Digital X-ray radiogrammetry measurement on a hand X-ray image

**Figure 2:** Image acquisition of a cadaver hand phantom with supination.
**Figure 3:** Effect of deviations from DXR image capture protocol on DXR-BMD. Symbols indicate phantom specimen.