Breast Composition Measurements Using Retrospective Standard Mammogram Form (SMF)

Ralph Highnam[1], Xia-Bo Pan[1], Ruth Warren[2], Mona Jeffreys[3], George Davey Smith[4], and Michael Brady[5]

[1] Siemens Molecular Imaging Ltd, Oxford, UK
[2] Consultant Radiologist, Addenbrooke's Hospital, Cambridge, UK
[3] Senior Lecturer in Public Health, Massey University, Wellington, NZ
[4] Professor of Clinical Epidemiology, University of Bristol, UK
[5] Professor of Information Engineering, University of Oxford, UK
rphighnam@aol.com

Abstract. Standard Mammogram Form (SMF), is a standardized, quantitative representation of a breast x-ray that can be easily estimated. From SMF it is possible to compute the volume of non-fat tissue and measures of breast density, both of which are of significant interest in determining breast cancer risk. Previous theoretical analysis of SMF suggested that a complete and substantial set of calibration data (such as mAs and kVp) would be needed to generate realistic breast composition measures, which is problematical since there have been many interesting trials that have retrospectively collected images with no calibration data. In this paper, we show how implementations of SMF include self-compensation mechanisms, so that SMF can be applied retrospectively to data for which calibration parameters are not (or only partially) available. To illustrate our findings, the current implementation of SMF (version 2.2β) was run over 4,028 digitized film-screen mammograms taken from 6 sites during the years 1988-2002, both with and without using the known calibration data. Results show that the SMF implementation running with no calibration data generates results which display a strong relationship with those obtained using a complete set of calibration data. More importantly, they bear a close relationship to an expert's visual assessment of breast composition using established techniques.

1 Background

The Standard Mammogram Form (SMF) representation of an x-ray mammogram is a standardized, quantitative representation of the breast (Highnam & Brady 1999) from which the volume and percentage of non-fat tissue can straightforwardly and automatically be estimated. Both the volume and percentage of dense tissue appear to be of significance for determining breast cancer risk (Boyd et al 1998, Heine and Malhotra 2002, Hufton et al 2004, Pawluczyk et al 2003). Recent work on SMF has shown that the estimate of percentage of non-fat tissue (SMF%) correlates strongly with an expert's visual assessment of breast density (Jeffreys et al. 2006), and SMF-based estimates of both volume (SMF Volume) and density show a small but

Susan M. Astley et al. (Eds.): IWDM 2006, LNCS 4046, pp. 243 – 250, 2006.

significant association with age as well as correlation with important known breast cancer risk measures such as body mass index (Jeffreys et al. 2003a).

We have previously reported (Highnam and Brady 1999) that implementations of SMF require a characterisation of the imaging system in the form of a set of calibration data, including parameters such as mAs, kVp and breast thickness. Furthermore, we presented a theoretical (Taylor's series) analysis that aimed to determine the errors in SMF values as a function of errors in the parameters, eg mAs. Unfortunately, and particularly when seeking to apply SMF to quantify films retrospectively, it is often the case that insufficient calibration data is available. In this paper, we present a new analysis of the implementations of SMF which show that they are in fact able to overcome both a lack of calibration data, and errors in the provided calibration data. The results contained in this paper are from SMF implementation 2.2β.

2 Calibration Parameter Compensation

The Taylor-series-based theoretical analysis of SMF by Highnam and Brady (1999) firmly conclude that both SMF% and SMF Volume are highly susceptible to errors in the calibration data. As a specific example, the analysis implied that a change in breast thickness of just 0.1cm changes SMF% by approximately 5%. However, documented evidence about errors in breast thickness readings from most mammography machines (Burch and Law 1995) and the lack of recorded breast thickness readings for many mammograms, led to an implementation of SMF which always estimates breast thickness directly from the image. The upshot is that, though apparently subtle, this implementation detail has a profound consequence: errors in the calibration data are used in the estimation of breast thickness, so the calibration data parameters are not independent, and, as a consequence, the Taylor's series analysis turns out to be massively overly pessimistic. Instead, we realise that the SMF process embodies a set of mutual constraints between parameter values, and these have the welcome property of automatically correcting for errors in the calibration data by the use of "ground truth" from the image. This constraint propagation process we call *Calibration Parameter Compensation (CPC).*

Following the approach of Tromans (2006), we illustrate CPC by considering SMF as a transfer function from input pixel value to thickness of non-fat tissue (h_{int}). Now consider that from the image itself, via the breast thickness estimation method, we know that a certain pixel value maps to $h_{int}=0$. It does not matter what the calibration data is, that mapping will remain constant and the breast thickness will be adjusted to keep it so. In short, the breast thickness is adjusted to compensate for any and all calibration data errors by using image-derived ground truth.

As an example, refer to Figure 1. The thick black line shows the "true" transfer function, that is, using the correct values of 61mAs and H=5.0cm. It also shows the transfer functions for the case where the mAs is deliberately made erroneous by a large amount: continuous thin line 40mAs (squares) and 100mAs (triangles) but breast thickness estimation (CPC) is not used. Finally, we show as dotted lines the resulting transfer functions when CPC is used. The legend notes the breast thicknesses estimated. Clearly, with CPC the transfer function is evidently far better