

**Hologic, Inc.**  
**Q&A Document – Oslo Tomosynthesis Screening Trial**  
**January 9, 2013**

**SUMMARY:**

The addition of 3D mammography screening significantly increased the number of invasive breast cancers found across all populations of women in the first large scale, peer-reviewed, clinical study comparing the addition of 3D (tomosynthesis) mammography to traditional 2D digital mammography alone. This study, published in *Radiology*, was designed to determine if 3D screening would find cancers that would be missed by 2D screening alone.

3D mammography was approved in the United States in February of 2011 for breast cancer screening and diagnosis and has been available in Europe and other regions of the world since 2008. Unlike a screening digital mammogram, which involves a single X-ray image of a breast, 3D mammography captures multiple, low-dose images from different angles around the breast. The images are then used to produce a 3D reconstruction of the breast.

From the technologist and patient perspective, there is little difference in the imaging time. Hologic 3D mammography is performed at the same time as a standard 2D mammography, uses the same machine, and takes just a few additional seconds.

Led by Per Skaane, M.D., at Oslo University Hospital Ullevaal, the study of more than 12,600 women showed a 40 percent increase for invasive cancers, along with a 15 percent decrease in false positives. Notably, this increase in invasive cancers detection was found across women of all breast tissue densities, both dense and fatty. At the same time, there was no increase in the detection of Ductal Carcinoma In Situ (DCIS), which is non-invasive and is often cited by critics of screening mammography as being over-diagnosed.

**BACKGROUND OF STUDY:**

**Q: How was this study designed and what were the primary goals for this study?**

A: The Oslo Tomosynthesis Screening Trial is the first large scale prospective breast cancer screening study investigating the use of tomosynthesis. The primary goal of this study was to determine if the cancer detection rate would increase with the addition of 3D. Unlike retrospective reader studies that do not affect clinical management, this study was actually incorporated into the clinical practice. Hence participants' clinical management was affected by decisions the radiologists made. In this study, each case was interpreted by four radiologists. Each of the four radiologists was presented with the following images: Arm A) 2D alone; Arm B) 2D+CAD; Arm C) 2D+3D; and Arm D) a synthesized 2D+3D. This study design allowed comparison of radiologists' performance with access to different images. This paper compared

two reading methods; namely 2D alone (Arm A) versus 2D+3D (Arm C). These methods of interpretation were predetermined as the primary reading methods for comparison.

**Q: Why were these two modes selected for the primary analysis?**

A: The 2D mode is currently the conventional method used for mammography screening. The 2D reader used standard digital mammography images for interpretation. The 2D+3D mode was recently approved for screening in the U.S. The study design allowed comparison of the current conventional method with the recently approved method.

**Q: Did this study compare single reading or double reading?**

A: The clinical practice in Oslo (and in many other parts of Europe) involves the independent reading of each case by two radiologists. If a case is suspicious for a finding that may be cancer, it is referred to an arbitration meeting where two or more radiologists review the case together with all available information about the case and make a decision whether or not to recall the woman for additional diagnostic procedures. At the first step, namely the independent reading, each of the radiologists typically refers to arbitration on average about 6 percent of cases. At arbitration many of these cases are dismissed as negative or benign, while some are actually recalled to come back for additional diagnostic procedures. The practice of having two independent readings of each case is called double reading. The meeting in which all suspected cases are reviewed and a clinical decision is rendered is called a consensus or an arbitration meeting.

In North America the conventional method of reading a mammogram is for a single radiologist to interpret the images and make a recommendation. This paper compared the interpretation of 2D images by a single radiologist (Arm A) to the interpretation of 2D+3D images by another radiologist (Arm C) and was therefore similar to the North American model of mammography interpretation. The primary endpoint was to compare cancer detection between these two arms of the study.

**Q: The study provides a comparison of false positive scores and actual recalls; how do these differ?**

A: The false positive scores are similar to a recall in North America in that the single reader interpreting the images called the examination “positive” (radiologist believed the images demonstrated a suspicious region needing further evaluation). A case was determined to be a false positive if a cancer was not found after further review or additional imaging. This study also reported actual recall rates after the arbitration meeting; however, the study design did not allow for an accurate comparison of actual recall rates. This was a result of the arbitration meeting design for this study. To provide the best care for patients, the arbitration meeting had access to all images (2D and 3D) regardless of which reading mode resulted in referring the case to arbitration as being suspicious. If a case was referred to arbitration by a reader with access to only 2D images, the case could be dismissed at arbitration because the radiologists also had

access to the 3D images. Thus the arbitration meeting was biased toward lowering recall rate for 2D alone (Arm A). Because of the bias related to the arbitration meeting, the false positive scores rather than the recall after arbitration is a better measure of radiologist performance.

## **KEY FINDINGS:**

### **Q: What are the key findings of this study?**

A: The key findings are summarized below. For 2D+3D compared to 2D alone:

- **A 40% increase in the detection of invasive breast cancers**
- A 27% increase in the detection of all cancers (invasive and in situ cancers combined)
- A 15% decrease in false-positive rates

All of these findings showed a statistically significant improvement. Unlike most other proposed breast cancer screening technologies, the addition of tomosynthesis resulted in improved cancer detection and simultaneously fewer false positives.

### **Q: How do these results compare to studies in the U.S.?**

A: At the Radiological Society of North America meeting in late November/early December of 2012, similar results of improved cancer detection were reported, with even larger reductions in recall rates from multiple institutions.

It is important to emphasize that the Oslo study is the first large-scale prospective clinical trial to be reported in a peer-reviewed journal.

### **Q: A 15% reduction in false positive scores at screening is reported in this study. Why have larger reductions in recall rates been reported in the U.S.?**

A: In Oslo an individual radiologist gives about 6% of cases a false positive score while in the U.S. on average about 10% of cases are recalled (about 95% of these cases are false positives). Because U.S. radiologists recall a much higher fraction of cases than radiologists in Oslo there is the potential for larger reductions in the U.S. compared to Oslo.

### **Q: Were the results of the study statistically significant?**

A: Yes. The study results showed a statistically significant improvement with regard to increase in cancer detection, increase in invasive cancer detection and decrease in false positive scores with 2D+3D compared to 2D alone.

**Q: Was the increase in cancer detection only for women with dense breasts?**

A: No, the addition of tomosynthesis resulted in increased cancer detection for women with both fatty and dense breasts.

**Q: Was there an increase in the detection of DCIS (Ductal Carcinoma In Situ)?**

A: No, there was no increase in the detection for DCIS. Many critics of mammography screening have suggested that mammography screening is finding too many DCIS cases and these cancers may contribute to overdiagnosis (detection of cancers that would not be fatal even if left untreated). All of the gain in detection with the addition of 3D was from additional invasive cancers detected.

**Q: What was the increase in reading time and what are the implications?**

A: The reading time for 3D mammography is more than the time to read digital, because there is more information for the radiologist to review (this is common and expected when a 3D technology augments a 2D technology). While the radiologist is spending additional time on the reading of screening exams, the resulting decrease in the number of diagnostic work ups improves overall practice workflow and decreases the amount of time required for interpreting diagnostic examinations.