*radiological interpretation,*
*ontology, diagnostic accuracy,*
*computer-aided diagnosis*

Teresa PODSIADŁY-MARCZYKOWSKA*,
Anna WRÓBLEWSKA**, Artur PRZELASKOWSKI**

# ENHANCEMENT OF ACCURACY IN MAMMOGRAMS INTERPRETATION USING ONTOLOGY-DRIVEN EDITOR FOR LESIONS DESCRIPTION AND CAD TOOL – PRELIMINARY RESULTS

Although mammography is the standard of reference for the detection of early breast cancer, as many as 25% of breast cancers may be missed. To reduce the possibility of missing a cancer, the following methods and tools has been proposed: continuing education and training, prospective double reading, retrospective evaluation of missed cases, and use of computer-aided detection (CAD). In the presented paper we report on preliminary results of reducing the number of false-negative cases in mammograms interpretation by using ontology-driven editor for mammograms description, and MammoViewer, a CAD tool for radiologists' perception improvement. The use of editor resulted in reduction of interpretation errors and improved consistency of diagnosis. Computerized image processing methods make the signs of pathologies more conspicuous and so resulted in improvement of lesion perception.

## 1. INTRODUCTION

Mammography is currently the most effective tool for early detection of breast cancer. It is still the only technique that can detect breast cancer in a preinvasive stage. Early detection via mammography increases breast cancer treatment options and the survival rate [1]. However, the diagnostic value of mammography is limited by significant and high, up to 25% rate of missed breast cancers [2,3]. Main, human-oriented causes lowering it are radiologists' perception and interpretation errors. In [6] authors state that in a practice where radiologists are well trained, experienced, and where a medical audit indicates satisfactory performance the frequency of missed malignant lesions indicates that such cases only rarely are effect of negligence. Therefore, strategies, means and tools to reduce radiologists' errors are essential to improve diagnostic effectiveness in mammography.

Methods described in medical literature and commonly used to decrease the number of missed cancers include: continuing education and training, prospective double reading, retrospective evaluation of missed cases, and use of computer-aided detection (CAD) [6]. Computer-aided detection (CAD) and ontology have been discussed among others in [9,10] as aiding tools for the interpretation of mammograms. In this paper we report on preliminary results of reducing the number of false-negative cases in mammograms

* Institute of Biocybernetics and Biomedical Engineering PAS, Trojdena 4, 02-109 Warszawa, Poland
** Institute of Radioelectronics, Nowowiejska 15/19, 00-665 Warszawa, Poland

interpretation by using ontology-driven editor for mammograms description, and MammoViewer, a CAD tool for radiologists' perception improvement. The use of editor resulted in reduction of interpretation errors and improved consistency of diagnosis. Computerized image processing methods make the signs of pathologies more conspicuous and so resulted in improvement of lesion perception.

The paper is organized as follows: the second section presents results of broad range review of medical literature searching to determine what are the causes of most commonly missed cancers, their types and features. The aim of the analysis was the selection of test sample composed of cases that will be probably false-negative when interpreting mammograms. Next section describes the tools proposed to reduce the false-negative rate. Fourth section contains description of test organization and test case samples followed by the results of interpretation with the use of editor and the results of perception improvement. Conclusions summarize received achievements.

## 2. RADIOLOGIST'S ERRORS, TYPES AND FEATURES OF MOST COMMONLY MISSED LESIONS, REASONS FOR THE MISS

Improvement of diagnostic value of mammography requires an understanding of false-negative cases and their characteristics. According to [5], main causes of missed cancers can be related to radiologist's errors and to poor technical conditions (the last not in the scope of this paper). Radiologist's errors can bee classified as interpretation and perception errors [5,8].

Perception error occurs when the lesion is included in the field of view and visible but is not recognized by the radiologist. The lesion may or may not have subtle features of malignancy. Possible causes for perception errors include dense parenchyma obscuring a lesion, and "happy eye syndrome" (not looking for additional lesions when one abnormality is seen).

The second cause of missed breast cancers related to radiologist error is incorrect lesion interpretation, which occurs when an abnormality with suspect features is observed and reported, but is misinterpreted as being definitely or at least probably benign.

According to [5,6,7,8] types of lesions most frequently misinterpreted or overlooked are:

1. Cluster of microcalcifications (small or hardly visible in dense breasts – perception and interpretation errors)
2. Lesions with benign appearance – masses and clusters of microcalcifications (interpretation error) [6,4]
3. Small masses ( perception error)
4. Asymmetric density (perception and interpretation errors)
5. Architectural distortion (perception error)

# 3. TOOLS PROPOSED TO REDUCE RADIOLOGIST'S ERRORS

## 3.1. ONTOLOGY-DRIVEN EDITOR

Our assumption was that using reliable domain knowledge representation to design and control mammographic data entry has the potential to improve their semantics (i.e. meaning and significance) and completeness. For that purpose mammographic ontology [14] has been used as a partial set of design assumption in developing graphical editor for mammograms description. Radiologists are familiar with images, so the choice of graphical, iconic data presentation model for mammography report editor has been obvious. Concepts necessary for mammographic lesions description has been represented by graphical primitives – icons (fig. 3.). Data about lesions are captured and displayed using this iconic representation. Sound, complete lesion definitions based on ontological model of the domain and graphical data representation result in improved data completeness and in turn in better interpretation of lesion's features and diagnosis. The editor has been described in more details in [10].

## 3.2. MAMMOVIEWER – COMPUTER-AIDED DIAGNOSIS TOOL

MammoViewer is a computer-aided diagnosis application [12,13]. It utilizes effective methods of presentation, processing, analysis and interpretation of images. It can be used as an advanced viewer dedicated to medical images - particularly mammograms (fig. 2.) – with a widespread imaging options like measuring structures, operating in regions of interests, defining parameters of a sliding window, setting the range of the shown pixel values etc. On the other hand as a scientific software package - MammoViewer includes a variety of methods to process medical images; in mammography processing in multiresolution wavelet domain (proved its particular usefulness) appeared particularly useful.

Mammogram preprocessing refers to enhance perception of pathologies in order to provide more conspicuous pathology signs for radiologists that interpret mammogram exams. This was achieved by:

- denoising - removing or weakening noise - information insignificant for diagnosis,
- local contrast enhancement - amplifying contours of lesions (which makes one can better determine shapes of abnormalities and evaluate their margins in more details) and emphasizing other diagnostically important lesion features like texture inside pathologies.

# 4. TESTS CASE SAMPLE

## 4.1. TEST SAMPLE FOR EDITOR TEST

For the editor test there were chosen the following types of cases (which has an impact on interpretation errors): clusters of microcalcifications (according to [4] radiologists evaluate clusters worse than masses), and cancerous lesions with benign appearance [6]. In the first part of the test cases were described using a medical image viewer (without scientific processing methods) and the reports were dictated, in the second part - both tools MammoViewer and editor were used for cases assessment.

Table 1. Lesions misdiagnosed in the first part of the test, included in the second part of the test. Average lesion subtlety – 3.1, average breast density – 2.67, last two masses obvious in fatty breasts.

| Case no. | Lesion Type | Breast Density | Subtlety |
|---|---|---|---|
| 1 | Mass | 2 | 1 |
| 2 | Micr. cluster | 4 | 2 |
| 3 | Micr. cluster | 4 | 3 |
| 4 | Micr. linear | 2 | 3 |
| 5 | Micr. cluster | 4 | 3 |
| 6 | Mass | 2 | 3 |
| 7 | Mass | 3 | 3 |
| 8 | Mass | 1 | 5 |
| 9 | Mass | 1 | 5 |

Table 2. Lesions with incorrect diagnostic process (different from DDSM) in the first part of the test, included in the second part of the test. Average breast density – 2.9, average lesion subtlety – 2.2

| Case no. | Lesion Type | Breast Dens. | Subt. |
|---|---|---|---|
| 1 | Mass | 3 | 3 |
| 2 | Mass | 3 | 1 |
| 3 | Micr.linear | 1 | 3 |
| 4 | Mass | 4 | 2 |
| 5 | Micr. cluster | 4 | 2 |
| 6 | Micr. cluster | 2 | 3 |
| 7 | Micr. cluster | 4 | 2 |
| 8 | Micr. cluster | 3 | 1 |
| 9 | Micr. cluster | 3 | 3 |
| 10 | Micr. linear | 3 | 3 |
| 11 | Mass | 2 | 3 |
| 12 | Micr. cluster | 3 | 2 |
| 13 | Mass | 4 | 2 |
| 14 | Mass | 2 | 1 |

### 4.2. PERCEPTION IMPROVEMENT TESTS

For the perception improvement test there were chosen the following types of cases: clusters of small microcalcifications in dense breast tissue, small masses, asymmetric densities. Preliminary tests were performed on a set of 16 mammograms from DDSM [11] (digitised at a pixel size of 43.5 microns and 50 microns with a 12-bit greyscale) containing pathologies – each lesion was shown in two projections (so there were 8 cases of lesions). Among the test images there were 7 cases (14 images) with spiculated masses: 2 cases (4 images) had microcalcifications within a mass, among the other 5 cases (without any microcalcifications) one case contained a mass with a partially ill-defined margin described in DDSM as a circumscribed mass. One case in the test set contained 6 benign circumscribed masses.

Within the tests one radiologist, expert in mammogram diagnosis, compared processed images with original ones and gave their opinion which was measured with subjective, comparative measures of quality by diagnostic symptoms analysis (tab. 3.).

Table 3. Subjective, comparative mark scale that was used in perception improvement tests. All important image features were scored consequently and conditions of true diagnosis were assessed.

| Mark scale | Wordy description of diagnostic image quality |
|---|---|
| +3 | definitely (arbitrarily) better |
| +2 | better |
| +1 | slightly better |
| 0 | comparable with the original |
| -1 | slightly worse |
| -2 | worse |
| -3 | definitely (arbitrarily) worse |

# 5. PRELIMINARY TEST RESULTS

## 5.1. PRELIMINARY ASSESSMENT OF DIAGNOSIS IMPROVEMENT

The use of editor in 23 cases of misdiagnosed pathologies resulted in diagnosis improvement. This general result should be divided into two categories:

- improvement when initial, fault diagnosis is benign or probably benign (9 cases)
- improvement when initial, incorrect diagnosis is additional exams needed, or suspicious (14 cases)

Change of BI-RADS diagnosis rating in the first category is more important because there is a substantial difference in the radiologist opinion and their consequences for the patient. In the second category the change is not so deep and can be described rather as improvement in diagnosis consistency, but leads to more appropriate diagnostic process.

The results for 9 cases of apparently benign cancerous lesions are shown in fig. 1. The diagnosis of the same case obtained in the first and second part of the test is presented and compared to DDSM assessment.

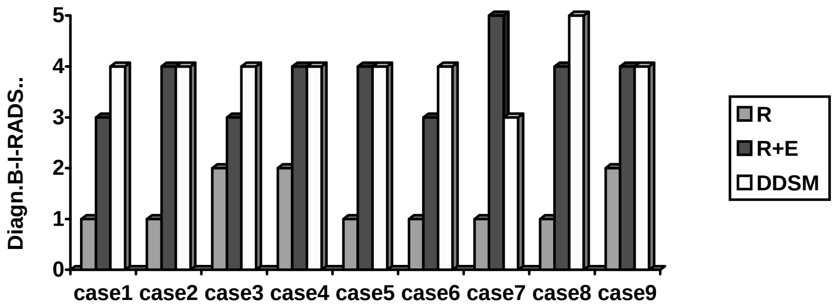


Fig.1. Effect of using graphical editor for correction of misdiagnosis - 9 cases of cancers misdiagnosed in the first part of the test as benign or probably benign – substantial correction in lesions diagnosis.
Figure legend: R – lesion diagnosis when radiologist uses CAD tool as medical viewer and report is dictated; R+E – lesion diagnosis when radiologist uses CAD tool as medical viewer and editor for lesion description and assessment; DDSM – lesion diagnosis as assessed in DDSM – a reference database.

An example of perfectly visible, but misinterpreted in the first part of the test lesion is presented and commented in Fig. 2. Fig. 3 presents description of the same case using editor and comments on diagnosis improvement.

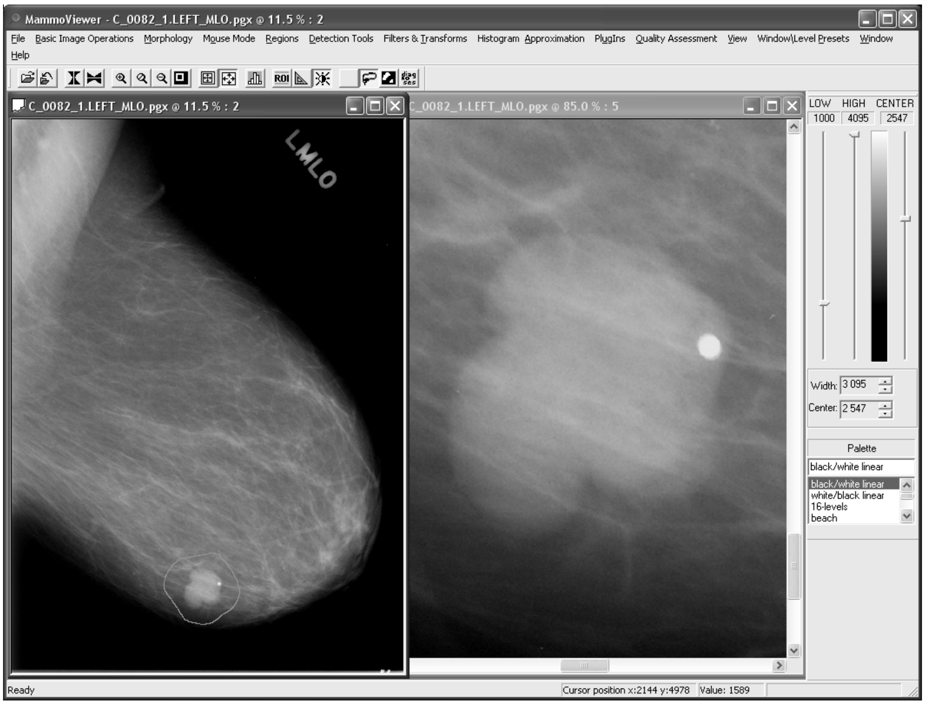


Fig.2. Case 8 – example of cancer with benign appearance. Apparently benign mass, perfectly visible in fatty breasts and additionally marked (left panel, gray outline). In the first round of the test, the mass has been described by radiologist in the dictated report as “lobulated, well-defined - benign“. The radiologist erroneously judged the lesion by its most benign features. However, the mass presents two suspected features: high density and partially ill-defined margin (right panel), not taken into account by radiologist in first test round. In fig. 3 (below) description of the same lesion using editor is presented.

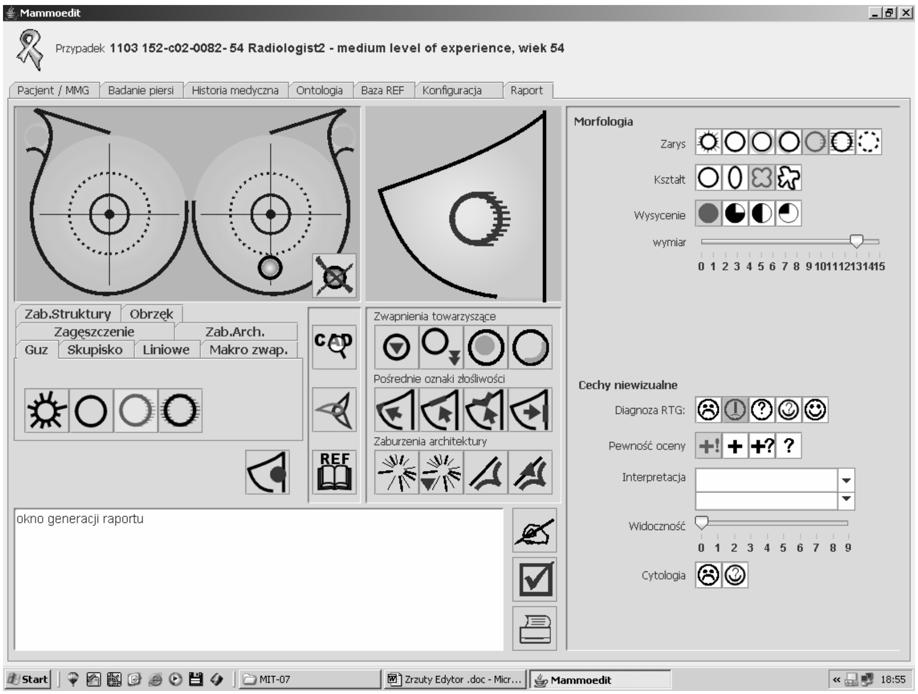


Fig.3. The mass with seemingly benign appearance (case 8) described using editor. Marked previously omitted suspected features: high density and partially ill-defined margin and in result mass diagnosis corrected. The use of editor enhance the completeness and accuracy of lesion features description, leading to better interpretation and decreasing the false-negative rate.

Improvement in diagnosis consistency - 14 cases – is presented in fig. 4.

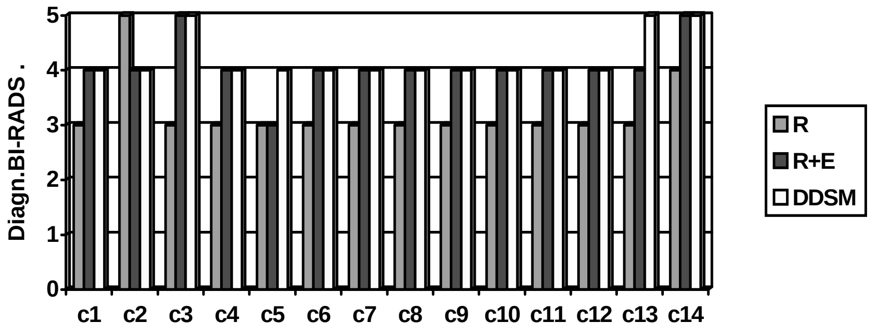


Fig.4. Effect of using graphical editor for improvement of diagnosis consistency. Figure legend is the same as in fig. 1.

5.2. PRELIMINARY ASSESSMENT OF PERCEPTION IMPROVEMENT

The obtained results are shown in tables 2 and 3. Exemplar pathologies - original and processed images are shown in Fig. 2. The results confirmed the effectiveness of the proposed methods for all types of pathologies with different subtlety.

Table 4. Mean mark of radiologist taking part in the tests.

| **Types of lesions** | **All types** | **Spicu-lated masses** | **Spiculated masses without microcal-cifications** | **Microcal-cifications within masses** | **Circum-scribed masses** |
|---|---|---|---|---|---|
| Number of images | 16 | 14 | 10 | 4 | 2 |
| Average subtlety (visibility of lesions: 1- the weakest lesions, 5 – the more obvious cases) | 3.4 | 3.21 | 3.3 | 3 | 5 |
| Mean radiologist's mark | **+2.5** | **+2.57** | **+2.6** | **+2.75** | **+2** |

Table 5. Mean mark of radiologist taking part in the tests.

| **Subtlety** | **1-2 (the weakest signs of pathology)** | **3** | **5 (the more obvious lesions)** |
|---|---|---|---|
| Number of images | 3 | 7 | 6 |
| Mean radiologist's mark | **+2** | **+2.71** | **+2.5** |

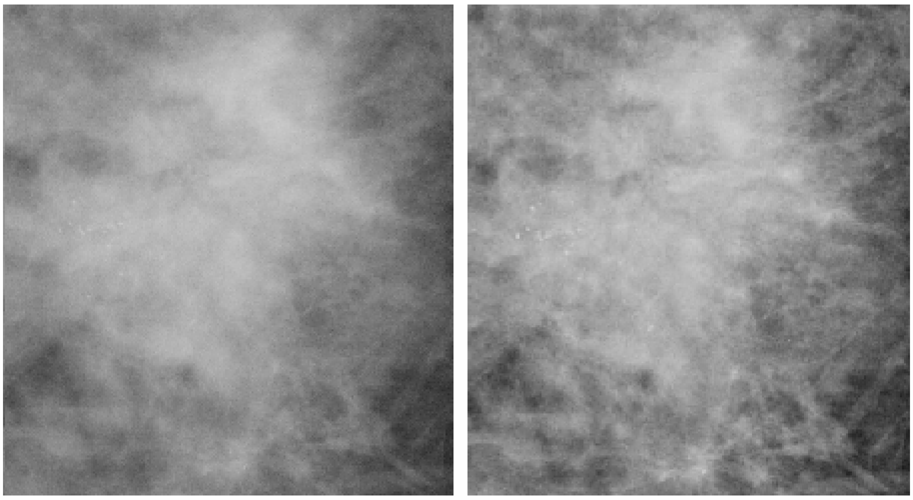

Fig.5. Example of perception improvement - spiculated mass with microcalcifications. On the left - original image and on the right - processed one. Diagnostically important lesion features – spiculated margin and inner microcalcifications – are better visible in the processed image.

## 6. CONCLUSIONS

While interpreting mammograms the use of the editor resulted in diagnosis improvement according to the preliminary results. There was the substantial change when initial, fault diagnosis was benign or probably benign (BI-RADS 1, 2 for 3,4 and 5) – the mistakes were previously made even for lesions that were very easy to capture (subtlety 5). The second effect was the improvement of diagnosis consistency where further diagnostic process was more appropriate.

For the perception improvement test the radiologist's marks showed the effectiveness of the used processing methods. The methods of perception improvement not only make the pathologies better visible but enhance their morphological features as well. Therefore, especially for inexperienced radiologist there is a need to use the both tools: an ontology-driven editor and a CAD tool.

In the tests a radiologist with an average level of experience made more mistakes for lesions particularly difficult to see and lesions with apparently benign morphology, whereas the beginning mammographer needed improvement of overall weak perception and interpretation. Therefore we are planning clinical tests with a greater number of cases and with a cooperation of 3 radiologists with different levels of expertise.

## 7. ACKNOWLEDGMENT

This work was supported by Foundation for the Development of Radiocommunication and Multimedia Technologies

BIBLIOGRAPHY

[1] LEE C. H., Screening mammography: proven benefit, continued controversy. Radiologic Clinics of North America, Vol. 40, pp. 395-407, 2002.
[2] SICKLES E., WOLVERTON D., DEE K., Performance parameters for screening and diagnostic mammography: specialist and general radiologists. Radiology, Vol. 224, pp.861-869, 2002.
[3] ELMORE J., et al., Screening mammograms by community radiologists: variability in false-positive rates. Journal of the National Cancer Institute, Vol. 94, pp. 1373-1380, 2002.
[4] MARKEY M. K., LO J.Y., FLOYD C.E., Masses vs calc differences between computer-aided diagnosis of breast masses and that of calcifications. Radiology, Vol. 223, pp. 489-493, 2002.
[5] MAJID A.S., DE PRADES E.S., DOHERTY R.D., SHARMA N.R., SALVADOR X., Missed breast carcinoma: pitfalls and pearls. RadioGraphics, Vol. 23, pp. 881-895, 2003.
[6] BIRD R.E., WALLANCE T.W., YANKASKAS B.C., Analysis of cancers missed at screening mammography. Radiology, Vol. 184, pp. 613-7, 1992.
[7] BIRDWELL R.L., IKEDA D.M., O'SHAUGHNESSY K.F., SISCKLES E.A., Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. Radiology, Vol. 219, pp. 192-202, 2001.
[8] MICHELL M.J., AL-ATTAR M.A, Detection of subtle mammographic sign of malignancy. Applied Radiology, Vol. 32, No. 9, pp. 11-16, 2003.
[9] PRZELASKOWSKI A., PODSIADŁY-MARCZYKOWSKA T., WRÓBLEWSKA A., BONIŃSKI P., BARGIEŁ P., Computer aided interpretation of medical images: limitations and challenges. Machine Graphics and Vision, in print, 2007.
[10] PODSIADŁY-MARCZYKOWSKA T., PRZELASKOWSKI A., WRÓBLEWSKA A., BONIŃSKI P., A framework for integrated, diagnosis supporting interface for mammograms description - advantages and pitfalls, Proc. Human-Computer Interaction IASTED-HCI, Chamonix, France, 2007.
[11] http://marathon.csee.usf.edu/Mammography/Database.html
[12] WRÓBLEWSKA A., PRZELASKOWSKI A., BARGIEŁ P., BONIŃSKI P., MammoViewer – CAD application based on effective multiscale image analysis. Polish Journal of Medical Physics & Engineering, Vol. 11, No. 1, pp. 17-28, 2005.
[13] http://www.ire.pw.edu.pl/MammoViewer
[14] PODSIADŁY-MARCZYKOWSKA T., GUZIK A., Mammographic ontology, model structure, definitions and concept instances. Bio-Algorithms and Medical Systems, Vol. 1, No. 1, pp. 247-252, 2005.