A Two-Phase Study Model for the Standardization of HER2 Immunohistochemical Assay on Invasive Ductal Carcinoma of the Breast

Pichet Sampatanukul MD, MSc, Benjaporn Chaiwun MD**, Sansanee Wongwaisayawan MD***, Parvinee Suwanagool MD****, Songkhun Vinyuvat MD*****; Anant Karalak MD******; Niphon Praditphol MD*******; Paisit Pauksakon MD********; Freecha Ruangvejvorachai BSc*, Andrew S Field FRCPA, FIAC, DipCyt(RCPA)*********, Pongsak Wannakrairot MD*

* Department of Pathology, Faculty of Medicine, Chulalongkorn University
** Department of Pathology, Faculty of Medicine, Chiang Mai University
*** Department of Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University
**** Department of Pathology, Faculty of Medicine, Siriraj Hospital, Mahidol University
***** Institute of Pathology, Department of Medical Service, Ministry of Public Health
****** Pathology Division, National Cancer Institute
******* Pathology Division, Rajavithi Hospital
******** Department of Pathology, Phramongkutklao College of Medicine
********* St Vincent’s Hospital, Sydney, Australia

Objectives: To develop and verify a standardized protocol for HER2 immunohistochemical assays on invasive ductal carcinoma of the breast in Thailand.

Material and Method: A two-phase study approach was employed. In the Phase One, after verifying the proposed protocol that adopted the HercepTest procedure using readily available primary antibodies, CB11 and A0485, Lab 1 performed the HER2 immunohistochemical staining for 137 cases of invasive ductal carcinoma twice with two types of the antibody. Nine pathologists from 8 centers independently examined and scored all the 2x137 stained slides that were blinded for antibody type. Interobserver reliability was calculated using pair-wise kappa. Following discussion of the results, the Phase Two study was planned. Lab 2 and Lab 3 independently performed the HER2 staining according to the protocol for 60 invasive breast carcinoma cases. The same group of pathologists scored 2x60 stained slides that were masked for laboratories. Interobserver reliability and interlaboratory agreement from each pathologist were calculated using kappa statistics. Three interpreted categories - namely negative, equivocal and positive tests were used in the analyses.

Results: Phase One study showed interobserver agreement between pairs varied from kappa 0.75 (95%CI, 0.68-0.82) to 0.06 (95%CI, 0-0.14) while Phase Two study obtained pair-wise kappa scores ranged from 0.84 (95%CI, 0.80-0.89) to 0.65 (95%CI, 0.59-0.71). Interlaboratory kappa for each pathologist was 0.67 (95%CI, 0.61-0.73).

Conclusion: The standardization of HER2 immunohistochemical assay was achieved through this two-phase study model. It had added benefits of improving pathologists’ expertise and verifying the HER2 testing protocol to be used in Thailand.

Keywords: HER2, IHC, Immunohistochemistry, Breast carcinoma, Guideline, Standardization

Full text. e-Journal: http://www.medassocthai.org/journal

Correspondence to : Sampatanukul P, Department of Pathology, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand. E-mail: fmedpst@md2.md.chula.ac.th
In Thailand, the incidence of breast cancer is on the rise and in Bangkok it is now the most common carcinoma in females\(^5\). The incidence in Bangkok was calculated at 25.4 per 100,000 and for the whole country was estimated at 17.2 per 100,000 during the period 1995-1997. Like many countries in this region, the emergence of the molecular-targeted chemotherapeutic agent, trastuzumab (Herceptin) has created the need for reliable HER2 testing. Furthermore, HER2 is one of the three bio-markers required for every newly diagnosed breast carcinoma according to the recommendation of the American Society of Clinical Oncology 2000\(^2\). At that time, no reference laboratory existed in Thailand, but seven large institutes provided immunohistochemistry (IHC) testing for HER2. Two primary antibodies were commercially available, the monoclonal antibody (Biogenex CB11 clone) and the polyclonal antibody (DAKO A0485). The relatively expensive HercepTest kit was not used. Fluorescent in situ hybridization for HER2 gene amplification testing was limited to research trials in a few centers. The pathologists’ experience was individualized and based on texts and journals with no centralized training courses or quality assurance measures in place.

According to an international review in 2003\(^3\), many countries have implemented national testing guidelines that vary in the level of detail and the number of recommendations. The concept is to use guidelines to standardize HER2 testing procedures at a local level. However, there are factors like laboratory experience and interobserver variability that influence the reliability of testing despite protocols.

To address the issue of establishing a reliable and reproducible IHC service for HER2 testing, a group of Thai pathologists that were working in breast pathology from 8 different institutes proposed a two-phase study to develop interobserver reliability and establish reproducibility between hospital laboratories.

The two-phase model was designed with a conceptual framework (Diagram 1). Phase One aimed to test interobserver reliability on reading HER2 IHC performed in one laboratory using a standardized protocol. Phase Two aimed to test interlaboratory agreement on IHC performed in each laboratory using the same protocol.

![Diagram 1. Conceptual Framework of the Reproducibility Test](image-url)
Material and Method

The proposed protocol - IHC method

Specimen: A block of tumor tissue should be cut and include benign breast tissue, and measure less than 15x15 mm and 4 mm in thickness.

Fixation: The tissue must be put into 10% neutral buffered formalin within an hour and be left until well-fixed but less than 48 hours.

Section: A freshly cut 3-micron thick section is recommended.

Block: Archive blocks kept at room temperature are not recommended for testing if they are more than two years old.

Antibody: Monoclonal antibody (clone CB-11, Biogenex), or Polyclonal antibody (A-0485, DAKO) with dilution at 1:500 and using 300 microlitre.

Staining procedure

1. Sections are incubated at 60 Celsius for 1 to 2 hours, deparaffinized and rehydrated with distilled water.
2. Antigen retrieval is performed with target retrieval solution (DAKO S1699) in a waterbath at 95-99 Celsius for 40 min.
3. The sections are then incubated at room temperature for 20 min, and rinsed in phosphate buffered saline for 3 min.
4. Endogenous peroxidase is blocked by 3% H2O2 in distilled water for 5 min.
5. Primary antibody incubation at room temperature for 30 min.
6. Visualization reagent (Envision, DAKO) incubation for 30 min.
7. Incubation with DAB for 10 min.

Reporting system

The invasive carcinoma component is assessed. The cell membrane staining, intensity of staining and the proportion of tumor cells staining are considered. The authors used three categories (1). “Negative” HER2 status includes the HercepTest scores 0 and 1+. It comprises absent or incomplete membrane staining of tumor cells or complete membrane staining in less than 10% of the assessed tumor cells. “Equivocal” HER2 status means the HercepTest score is 2+. It requires at least 10% tumor cells with complete but weak membrane staining. Included in this category are also those stains that do not meet satisfactory quality criteria but show some degree of membrane or cytoplasmic staining. “Positive” HER2 status is equal to the HercepTest score 3+, that is, at least 10% tumor cells show strong cytoplasmic membrane staining with satisfactory quality.

Quality satisfaction

Three parameters are used for the quality determination. Firstly, normal ductal epithelium has to be negative. Secondly, the positively stained tumor cells have to maintain nuclear integrity. And thirdly, the positively stained tumor cells should show malignant criteria, with malignant nuclei and intact cytoplasm.

The study scheme

In Phase One, Laboratory 1 performed the staining according to the proposed protocol with two antibodies for each case, that is, monoclonal antibody (clone CB11) and polyclonal antibody (A0485). Then nine pathologists (A, B, C, D, E, F, G, H and I) independently interpreted the results to assess interobserver reliability, and whether the antibody type effected the interobserver agreement. In Phase Two, Laboratories 2 and 3 performed the staining following the standardized protocol the antibody had shown to give better interobserver reliability. The same group of nine pathologists independently interpreted the results, to assess the reproducibility of staining results between Laboratories 2 and 3.

The pathologists

The nine pathologists practiced in seven different institutes and varied in their experience in reading and interpreting HER2 IHC, and in the number seen each month from a few to twenty.

The cases

Recent cases from the seven participating institutes were used, each containing invasive ductal carcinoma and benign breast elements. The fixation time was well-controlled and limited to less than 48 hours in 10% neutral buffered formalin. A total of 137 blocks was pooled, mixed and masked.

In Phase One, two sections were cut from each block, one for the monoclonal antibody and the other for polyclonal antibody. The staining was conducted in one laboratory (Laboratory 1). A total of 274 stained slides were blinded for antibody type and then independently scored by the nine participating pathologists according to the established criteria.

In Phase Two, 60 cases were randomly picked from the 137 cases in Phase One, and then the two separate Laboratories, 2 and 3, performed the IHC using the monoclonal antibody (Biogenex CB11 clone) that
had been chosen after the Phase One evaluation. A total of 120 slides were scored independently by the same pathologists, blinded to laboratory source.

Protocol development

In developing the protocol, a pilot study had been carried out to validate the result of staining with the CB11 clone and the A0485 antibodies against the HercepTest™. The authors noted that the preservation of nuclear detail in the stained slides was helpful in assessing whether excessive thermal effect a potential source of over-staining (Fig. 1). Therefore, the authors used this parameter of well-preserved nuclear features as an internal control for a satisfactory stain, in addition to the generally accepted parameters of an absence of stain in normal ductal epithelium and malignant morphology in the HER2-positive tumor cells.

The pathologist performances

The pathologists were gathered in a room containing multiple microscopes to look at all the stained slides and score each case independently. Phase One was scheduled on the 4-5 January 2003. Phase Two was scheduled on the 17-18 May 2003. There was an evaluation session of Phase One results between these two dates.

Data analysis and statistics

Data were analyzed to show the interpretation concordance between pairs of participants by kappa statistics. The four scoring categories of the HercepTest were collapsed to three interpretative categories of negative (0, 1+), equivocal (2+) and positive (3+). Each stained slide had a consensus result which derived from the majority score. The stained slides without majority scores were marked “no consensus”. The percentage of cases where at least seven pathologists agreed was calculated and used to compare two separate events. In Phase Two, an interlaboratory kappa was calculated according to individual interpretations of the coupled stained slides between the two different laboratories. A p-value of less than 0.05 was considered statistical significance.

Results

Phase One: A total of 274 slides were assessed. The breakdown of the number of cases falling into each score as reported by the nine pathologists is presented in Table 1. The consensus score, derived from the majority of pathologists in each case, showed HER2 3+ status in 24% and 2+ in 18% of cases. Interobserver agreement between pairs varied from kappa 0.75 (95%CI, 0.68-0.82) to 0.06 (95%CI, 0-0.14) (Table 2). Cases in which at least seven pathologists agreed showed concordance of 68.2%. There was no statistically significant difference between the agreement results of CB11 clone and A-0485 antibodies, though the former gave slightly more agreement cases in each category from complete agreement (all nine pathologists) to six or more agreements (Table 3).

Phase Two: A total of 120 slides were assessed independently by the same group of nine pathologists. Pair-wise kappa scores ranged from 0.84 (95%CI, 0.80-0.89) to 0.65 (95%CI, 0.59-0.71) (Table 4). At least seven observers concordance was achieved in 84% of the cases and there was no statistically significant difference.

Fig. 1  Demonstration of effect of over-heating on nuclear features. A. The appropriate temperature shows invasive tumor cells having intact nuclei. The IHC result is negative. B. The over-heating causes nuclear burn-out. The IHC result is falsely positive (Immunohistochemical staining for HER2)
Discussion

Although new technologies for HER2 testing may emerge\(^5\), IHC for HER2 uses familiar techniques and is a cost-effective test and will remain in routine clinical practice for HER2 screening in breast carcinoma. Many countries have national testing guidelines that vary slightly in specifications\(^6-10\), and issues related to fixation, the type of antibody, staining protocols and interobserver interpretation.

Fixation by various fixatives has been shown to impact on HER2 results\(^11, 12\). However, formalin fixation is the most commonly used and IHC is standardized on formalin fixation. Fixation time may influence IHC too. The HERA Trial recommends 18-24 hours in neutral buffered formalin as standard fixation\(^13\).

A number of commercially available antibodies to HER2 are on the market, some in kit forms offering greater potential reproducibility when protocols are strictly followed. These include the first FDA approved kit, the HercepTest. Its high price has reduced its use, however. The authors developed a protocol

---

**Table 1. Phase One Study: Pathologists’ scores of HER2 status**

<table>
<thead>
<tr>
<th>Pathologists</th>
<th>HER2 Negative (0/1+)</th>
<th>HER2equivocal (2+)</th>
<th>HER2 Positive (3+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>148</td>
<td>57</td>
<td>69</td>
</tr>
<tr>
<td>B</td>
<td>145</td>
<td>76</td>
<td>53</td>
</tr>
<tr>
<td>C</td>
<td>137</td>
<td>64</td>
<td>73</td>
</tr>
<tr>
<td>D</td>
<td>133</td>
<td>53</td>
<td>88</td>
</tr>
<tr>
<td>E</td>
<td>162</td>
<td>53</td>
<td>59</td>
</tr>
<tr>
<td>F</td>
<td>174</td>
<td>65</td>
<td>35</td>
</tr>
<tr>
<td>G</td>
<td>193</td>
<td>66</td>
<td>15</td>
</tr>
<tr>
<td>H</td>
<td>69</td>
<td>117</td>
<td>88</td>
</tr>
<tr>
<td>I</td>
<td>128</td>
<td>71</td>
<td>75</td>
</tr>
<tr>
<td>Consensus(^*) (%)</td>
<td>147 (53.6)</td>
<td>51 (18.6)</td>
<td>66 (24.1)</td>
</tr>
<tr>
<td>No consensus(^**) (%)</td>
<td>10 (3.6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^*\) Consensus = the majority score of individual slide; \(^**\) No consensus = slide that had no majority score

---

**Table 2. Phase One Study: Inter-observers’ kappa statistical analysis**

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.73</td>
<td>0.71</td>
<td>0.63</td>
<td>0.68</td>
<td>0.54</td>
<td>0.33</td>
<td>0.45</td>
<td>0.66</td>
</tr>
<tr>
<td>B</td>
<td>0.68</td>
<td>0.60</td>
<td>0.75</td>
<td>0.66</td>
<td>0.39</td>
<td>0.40</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.65</td>
<td>0.61</td>
<td>0.45</td>
<td>0.25</td>
<td>0.48</td>
<td>0.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.55</td>
<td>0.40</td>
<td>0.23</td>
<td>0.53</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>0.66</td>
<td>0.45</td>
<td>0.33</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>0.52</td>
<td>0.21</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>0.06</td>
<td>0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

between the two laboratories performing the staining with the CB 11 antibody (Table 5), although the complete agreement, at least 8 pathologists agreement and at least 7 pathologists agreement appeared slightly higher in Laboratory 3. The inter-laboratory kappa for each pathologist was 0.67 (95%CI, 0.61-0.73).

---

**Table 3. Phase One Study: Comparison of number of agreements between A0485 and CB11 clone antibodies**

<table>
<thead>
<tr>
<th>Number of agreements</th>
<th>A0485 N = 137</th>
<th>CB11 N =137</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree = 9</td>
<td>31 (22.6%)</td>
<td>39 (28.5%)</td>
<td>ns</td>
</tr>
<tr>
<td>Agree ≥ 8</td>
<td>69 (50.4%)</td>
<td>71 (51.8%)</td>
<td>ns</td>
</tr>
<tr>
<td>Agree ≥ 7</td>
<td>90 (65.7%)</td>
<td>97 (70.8%)</td>
<td>ns</td>
</tr>
<tr>
<td>Agree ≥ 6</td>
<td>109 (79.6%)</td>
<td>114 (83.2%)</td>
<td>ns</td>
</tr>
<tr>
<td>Agree ≥ 5</td>
<td>132 (96.4%)</td>
<td>132 (96.4%)</td>
<td>ns</td>
</tr>
<tr>
<td>No majority</td>
<td>5 (3.6%)</td>
<td>5 (3.6%)</td>
<td></td>
</tr>
</tbody>
</table>

ns = not significant
following the procedural standardized protocol of the HercepTest but using separately purchased primary antibodies and other components, which reduced the cost by six times. During this development, the authors observed that in some protocols, the tumor nuclei were destroyed due to an overheating effect. The authors speculated that this could be the source of false positive staining, and suggest that a temperature-controlled water-bath is more reliable than retrieval by microwave. In addition, benign epithelial cells adjacent the carcinoma act as a negative control. The authors established these two parameters as our requirements for satisfactory staining and our protocol has proved effective and valid in the large or central laboratories involved in the present study.

There is a potential source of error due to varying sensitivities of the HER2 antibodies[14]. The authors tested the two most commonly used antibodies by comparing the interobserver agreement to see if one gave more reliable results using our proposed method of staining. The results showed that there was no statistically significant difference.

Interobserver reliability depends not only on laboratory quality controls and the published criteria of scoring, but also fundamentally on the perception and experience of the pathologists. A previous study using the HercepTest showed an interobserver kappa statistic between well-trained pathologists of 0.4[15]. Another study using the HercepTest and five pathologists achieved complete agreement in 48%[16]. In the present study, the pathologists scored the slides independently and then met and discussed the cases together. In the first phase, it was evident that pathologists G and H interpreted the scoring differently from the majority of the group. The former scored too few cases as 3+, while the latter scored 2+ too frequently, introducing a major source of discrepancy in paired kappa statistics. In the Phase Two study, the training process and added experience in scoring produced a definite improvement in the interobserver reliability with a paired kappa statistic of 0.65 to 0.84, and complete agreement among the “nine raters” of nearly 70%.

A major problem in the Phase One study was the large number of cases: 137 cases were gathered from most of the participating centers in order to get a variety of sources. Each pathologist had to score a total of 274 slides (137 each for CB11 and A0485) in a two-day period, which was a heavy workload for pathologists and laboratories. In Phase Two, the number of cases was cut to 60, and the CB11 clone antibody was chosen because it had a slightly better, but not statistically significant, difference rate of agreement than the A0485 antibody. It could be argued that the decreased number of cases and workload lead to improve scoring concordance, but the degree of improvement in kappa concordance was very substantial.

The Phase One study assessed and inherently improved interobserver concordance in slide interpretation and scoring. The Phase Two study simultaneously assessed interobserver and interlaboratory concordance. The two laboratories showed high concordance of more than 80%, with a kappa statistic of interlaboratory reliability by each pathologist of 0.67. The present study emphasizes the need to train pathologists interpreting the staining and to ensure the laboratory quality control. There was a tendency for the local, small laboratories to over diagnose HER2 IHC 3+ cases[17]. The authors observed that over-heating can disrupt nuclear integrity and results in altered staining. The use of benign breast as an internal negative control has been reported[18], as has the observance that high-nuclear grade morphology is associated with positive HER2 status[19]. The authors, therefore, used these three features in our protocol as our satisfactory quality criteria.

### Table 4. Phase Two Study: Inter-observers’ kappa statistical analysis

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.73</td>
<td>0.80</td>
<td>0.78</td>
<td>0.78</td>
<td>0.68</td>
<td>0.69</td>
<td>0.84</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.67</td>
<td>0.75</td>
<td>0.74</td>
<td>0.71</td>
<td>0.76</td>
<td>0.81</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.77</td>
<td>0.67</td>
<td>0.65</td>
<td>0.72</td>
<td>0.83</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.74</td>
<td>0.68</td>
<td>0.75</td>
<td>0.83</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>0.78</td>
<td>0.77</td>
<td>0.82</td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>0.81</td>
<td>0.74</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>0.80</td>
<td>0.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.82</td>
</tr>
</tbody>
</table>

### Table 5. Phase Two Study: Comparison of number of agreements between Laboratories 2 and 3

<table>
<thead>
<tr>
<th>Number of agreements</th>
<th>Lab 2</th>
<th>Lab 3</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree = 9</td>
<td>39 (65.0%)</td>
<td>42 (70%)</td>
<td>ns</td>
</tr>
<tr>
<td>Agree ≥ 8</td>
<td>45 (75.0%)</td>
<td>48 (80%)</td>
<td>ns</td>
</tr>
<tr>
<td>Agree ≥ 7</td>
<td>50 (83.3%)</td>
<td>51 (85%)</td>
<td>ns</td>
</tr>
</tbody>
</table>

ns = not significant
As Schnitt SJ and Jacobs TW stated “pathologists are caught between a rock and a hard place” in the current approach to HER2 testing(19). It appears more logical to determine HER2 protein expression by IHC rather than the level of gene amplification by FISH when treatments such as Herceptin are specifically targeted toward the HER2 protein expressed on the cell surface. IHC analysis should be the more biologically relevant assay, particularly in cases in which there is discordance between gene copy number and the level of protein expression.

Therefore, the standardization of the IHC assay protocol is a crucial requirement, along with high standards of interobserver and interlaboratory concordance, which are essential on a national scale. The authors agree with the editorial comment on the guidelines for HER2 testing in the UK, that although an annual caseload volume seems to be related to accuracy, it does not ensure an accurate result(20). The important step is to decide how to enable laboratories to perform quality testing and pathologists to reproducibly score the staining. In our study, the two-phase model has proved a useful in achieving of these aims.

In conclusion, this two-phase study approach not only assessed interobserver concordance in interpretation and scoring of IHC for HER2, and interlaboratory concordance in performing the HER2 IHC. It had the added benefits of improving pathologists’ expertise and, most importantly, developing and verifying a protocol, using readily available and inexpensive primary antibodies, A0485 or CB11, that can be used throughout the region of Thailand as a national protocol for HER2 testing.

Acknowledgement

The authors wish to thank Roche Thailand who financed the antibodies and all facilities.

References
15. Tsuda H, Sasano H, Akiyama F, Kurosumi M,


รูปแบบการศึกษาแบบ 2 ระยะสำหรับการจัดทำมาตรฐานของการตรวจอิมมูโนเคมีของโปรตีน HER2 ในเนื้อเยื่อมะเร็งเต้านม

พิเชฐ สัมปทานุกุล, แบงก์พร ไขรัศริน, ศัณณีย์ วงศ์ไวศยวรรณ, ภัสสร วิทกุลภูพิม, ทรงศักดิ์ ภูพิม, ณัฐเดช กลักษณ์, นิพนธ์ ประดิษฐผล, ไพสิฐ เล็กสกนธ, ปริชา เรืองวชิรชัย, Andrew S Field, พงษ์ศักดิ์ วรรณไกรโรจน์

วัตถุประสงค์: เพื่อพัฒนาและพิสูจน์วิธีการที่จะใช้เป็นมาตรฐานของการทดสอบอิมมูโนเคมีของโปรตีน HER2 ในเนื้อเยื่อมะเร็งเต้านมในประเทศไทย

วัสดุและวิธีการ: ใช้รูปแบบการศึกษาแบบ 2 ระยะ ระยะที่ 1 ภายหลังจากได้ตรวจสอบการย้อมของวิธีการที่นำเสนอ โดยใช้ตามแบบวิธีการของ HercepTest และแอนติบอดีหลัก (CB11 และ A0485) ที่ซื้อแยกต่างหาก ทรงศักดิ์ ปฏิบัติการ ที่ 1 ทำการย้อมอิมมูโนเคมีของ HER2 ในชิ้นเนื้อมะเร็งจำนวน 137 ราย สองครั้ง ด้วยแอนติบอดี 2 ชนิด พร้อมที่จะย้อม 9 คน จาก 8 สถาบัน ตรวจจุดและแปลงแปลง สไลด์ย้อมจำนวน 2 x 137 แผ่น ซึ่งเป็นชนิดของแอนติบอดีโดยจัดระเบียบการประเมินความตรงกันของผลการย้อม ใช้การวิเคราะห์ kappa ของแต่ละกลุ่มของพยาธิแพทย์ เมื่อเก็บรายละเอียดผลแล้ว การศึกษาระยะที่ 2 โดยนำไปสนใจธีม ทรงศักดิ์ปฏิบัติการที่ 2 และ ที่ 3 ดำเนินการ HER2 ด้วยเทคนิคเดิมโดยมอบให้แพทย์ที่ทำการย้อมจำนวน 60 ราย พร้อมที่จะย้อมอิมมูโนเคมีตามที่ได้ทำการย้อมจำนวน 2 x 60 สไลด์ที่ทำการย้อมจากส่งเสริมกัน โดยเก็บรายละเอียดผลการจัดระเบียบการประเมินความตรงกันของผลการย้อมระหว่างสไลด์ พร้อมที่จะย้อม 70 ราย พร้อมที่จะปิดบังห้องปฏิบัติการที่ทำการย้อมสไลด์ โดยให้ค่า kappa ใช้การวิเคราะห์ kappa อีก 1 การวิเคราะห์ค่า kappa สำหรับการย้อม HER2 ด้วยเทคนิคชั้นสูงที่มีค่า kappa ค่าสูงสุดเป็น 0.75 95% CI 0.68-0.82 และค่า kappa ต่ำสุดเป็น 0.65 95% CI 0.59-0.71

ผลการศึกษา: การศึกษาระยะที่ 1 ความตรงกันของผลการเรียกผลลัพธ์ของพยาธิแพทย์ โดยค่า kappa สูงสุดคือ 0.75 95% CI 0.68-0.82 และค่า kappa ต่ำสุดคือ 0.65 95% CI 0.59-0.71

สรุป: การศึกษาได้รับมาตรฐานการย้อม HER2 ด้วยเทคนิคชั้นสูงที่มีค่า kappa สูงสุดถึง 0.75 95% CI 0.68-0.82 โดยใช้รูปแบบการศึกษาแบบ 2 ระยะ ระยะที่ 1 กำหนดด้วยการศึกษาชั้นสูงที่มีค่า kappa สูงสุดถึง 0.75 95% CI 0.68-0.82 สามารถนำไปใช้ได้ในประเทศไทย

1688 J Med Assoc Thai Vol. 88 No. 11 2005