



# Concordance assessment of Watson for Oncology in breast cancer chemotherapy: first China experience

Hong Pan<sup>1#</sup>, Jing Tao<sup>1,2#</sup>, Mengjia Qian<sup>1#</sup>, Wenbin Zhou<sup>1</sup>, Yi Qian<sup>1</sup>, Hui Xie<sup>1</sup>, Shengqi Jing<sup>3,4</sup>, Tingyu Xu<sup>3,4</sup>, Xin Zhang<sup>3,4</sup>, Zuolei Dai<sup>3,4</sup>, Mingliang You<sup>5</sup>, Yun Liu<sup>3,4</sup>, Xiaoan Liu<sup>1</sup>, Shui Wang<sup>1</sup>

<sup>1</sup>Department of Breast Surgery, The First Affiliated Hospital with Nanjing Medical University, Nanjing 210029, China; <sup>2</sup>Department of General Surgery, The Fourth Affiliated Hospital of Nanjing Medical University (Nanjing Pukou Hospital), Nanjing 210031, China; <sup>3</sup>Department of Information, The First Affiliated Hospital with Nanjing Medical University, Nanjing 210029, China; <sup>4</sup>Institute of Medical Informatics and Management, Nanjing Medical University, Nanjing 210029, China; <sup>5</sup>Hangzhou Cognitive Care, Hangzhou Artificial Intelligence Industrial Park, Hangzhou 310000, China

**Contributions:** (I) Conception and design: H Pan, Y Liu, X Liu, S Wang; (II) Administrative support: S Jing, M You; (III) Provision of study materials or patients: Y Qian, H Xie, T Xu, X Zhang; (IV) Collection and assembly of data: M Qian, W Zhou, S Jing; (V) Data analysis and interpretation: H Pan, J Tao, M Qian, W Zhou, Z Dai; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

<sup>#</sup>These authors contributed equally to this work.

**Correspondence to:** Dr. Yun Liu. Department of Information, The First Affiliated Hospital with Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, China. Email: liuyun@njmu.edu.cn; Dr. Xiaoan Liu. Department of Breast Surgery, The First Affiliated Hospital with Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, China. Email: liuxiaoan@126.com; Prof. Dr. Shui Wang. Department of Breast Surgery, The First Affiliated Hospital with Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, China. Email: ws0801@hotmail.com.

**Background:** IBM Watson for Oncology (WFO) is an artificial intelligence cognitive computing system that provides confidence-ranked, evidence-based treatment recommendations for cancer. We examine the level of agreement for breast cancer chemotherapy between WFO recommended and clinical use in a large population of breast cancer cases.

**Methods:** A total of 1,301 breast cancer patients were reviewed in The First Affiliated Hospital with Nanjing Medical University, China from June 2013 to December 2017. Patients' data were entered manually into WFO by the trained senior oncology fellows. Chemotherapy recommendations were provided in 3 categories, "Recommended", "For Consideration", and "Not Recommended". Concordance was achieved when oncologists' treatment suggestions were in the "Recommended" or "For Consideration" categories.

**Results:** The chemotherapy regimen concordance was 69.4% among all breast cancer cases, 65.0% among the cases in adjuvant chemotherapy (AC) group and 96.7% among the cases in neoadjuvant chemotherapy (NAC) group. The concordance varied greatly in subset analysis with respect to TNM stage and molecular subtype. AC recommendations were concordant in 92.3% of stage III breast cancer and 50.8% of stage I. However, the concordance varied by molecular subtype, which was higher for triple negative breast cancer (89.3%) than others. The chemotherapy regimen concordance declined significantly with increasing age, except for the age group 41–50 years.

**Conclusions:** Chemotherapy regimens provided by WFO did not exhibit a high degree of agreement with those suggested by oncologists in clinical practice in the hospital in China. The current effort is underway to enhance WFO's capabilities as a cognitive decision support tool by incorporating regional guidelines, enabling oncologists and patients to benefit from WFO worldwide.

**Keywords:** Watson for Oncology (WFO); breast cancer; chemotherapy; concordance

Submitted Sep 08, 2018. Accepted for publication Jan 04, 2019.

doi: 10.21037/tcr.2019.01.34

View this article at: <http://dx.doi.org/10.21037/tcr.2019.01.34>

## Introduction

Breast cancer is the most common cancer and the leading cause of cancer death among women worldwide (1). The widespread use of adjuvant systemic therapies, including chemotherapy, endocrine therapy, and anti-HER2 therapy, has contributed significantly to improving breast cancer outcomes (2-5).

Chemotherapy plays an important role in breast cancer systemic treatments. It was initially prescribed to patients using an adjuvant alkylating agent, based on the first randomized trial NSABP B-01 reported in 1968 (6). In contrast with surgery that evolved from more aggressive to less aggressive, chemotherapy regimens and their indications expanded (7). Adjuvant cytotoxic chemotherapy regimens have evolved from single alkylating agents to polychemotherapy regimens incorporating anthracyclines and/or taxanes. The widespread adoption of more effective chemotherapy regimens contributed to declining breast cancer mortality globally, however, it also resulted in many patients being unintentionally over-treated with chemotherapy who might be cured without chemotherapy (8,9). Nowadays, women with early-stage breast cancer are faced with challenging chemotherapy regimen decisions. Identifying which patients derive greatest benefit from chemotherapy and which regimen is best for them have been being discussed over the past decades (9,10).

Along with the growth of massive genetic and clinical databases, breast cancer treatment developed quickly, the cycle time for changes to breast cancer treatment guidelines was shorter and shorter. Currently, predictive factors that identify benefit from chemotherapy include classical clinicopathologic characteristics (tumor size and number of positive axillary lymph nodes), and expression of estrogen receptor (ER), progesterone receptor (PR), HER2 and Ki-67, as well as multiparameter gene expression assays (11-14). Maybe some other more effective predictive factors will be identified to realize fully the promise of precision chemotherapy in the near future. Oncologists who treat breast cancer are challenged by a large and rapidly expanding knowledge base. There is always little time available for oncologists to track and access relevant information using in the clinical practice (15). We have anticipated that an oncologist equipped with a computer-based decision support system will be able to select for any given patient an optimal regimen that maximizes therapeutic efficacy while minimizing the side effects correlating with ineffective treatments.

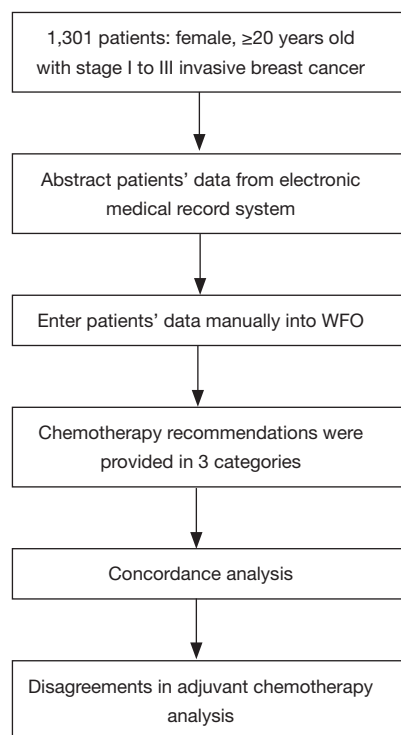
IBM Watson for Oncology (WFO) is an artificial intelligence cognitive computing system that provides confidence-ranked, evidence-based treatment recommendations for cancer. It was developed in collaboration with cancer experts at Memorial Sloan Kettering Cancer Center (MSKCC). The system provides oncologists with treatment options, which are derived from established guidelines, medical literature, and training from patient cases. A retrospective, observational study carried out in India showed that treatment decision made by WFO exhibited a high degree of agreement with those of the multidisciplinary tumor board (16). Based on their study, non-concordance may contribute to the demographic characteristics, such as comorbidity burden, patients' preferences, and level of social support systems, as well as different treatment guidelines.

In the present study, we examine the level of agreement for primary breast cancer chemotherapy [including adjuvant chemotherapy (AC) and neoadjuvant chemotherapy (NAC)] between WFO recommended and clinical use from The First Affiliated Hospital with Nanjing Medical University (Nanjing, China) in a large population of breast cancer cases. The study was aim to determine the chemotherapy concordance, and discuss the potential reasons of non-concordance.

## Methods

### *Patients and study design*

This retrospective, single-center, non-interventional study reviewed 1,301 breast cancer patients in Jiangsu Province Hospital with Nanjing Medical University, China from June 2013 to December 2017. The eligible patients were female,  $\geq 20$  years old with stage I to III invasive breast cancer, who received chemotherapy determined according to established clinical risk factors and breast cancer guidelines. Patients were excluded if they were male, had stage IV breast cancer, or relapsed after primary therapy. All the entered patients were separated into two groups: AC group included 1,121 patients who received AC after surgery, and NAC group included 180 patients who received NAC after diagnosis by biopsy. Oncologists determined all the chemotherapy regimens used in clinic. Three trained senior oncology fellows reviewed the electronic medical record system, and patients' clinical information and recommended treatment regimens between 2013 and 2017 were recorded. The design of the study is shown in *Figure 1*.



**Figure 1** The flow diagram of the study. WFO, Watson for Oncology.

The study protocol was approved by the ethics committee of The First Affiliated Hospital with Nanjing Medical University (2018-SR-368) and was conducted in accordance with the Helsinki Declaration of 1964 (revised 2008). Due to the retrospective nature of this study, the need for informed consent was waived.

### ***Pathologic assessment***

In the present study, patients diagnosed with metastasis by imaging examination were all excluded from our study. For patients in the AC group, pathological tumor size and axillary lymph node status were assessed after surgery, to determine pathologic TNM stage according to the American Joint Committee on Cancer (AJCC). For patients in the NAC group, clinical tumor size and clinical axillary lymph node status were assessed at diagnosis, to determine clinical TNM stage. Additionally, ER, PR, HER2 status, and Ki-67 were assessed from the biopsies taken at diagnosis or tumor tissues taken at surgery, to allocate tumors in molecular subtype classification, as suggested by the St Gallen 2017 consensus. The following classifications were used:

Luminal A: ER positive and PR positive (cut-point  $\geq 20\%$ ), HER2 negative, Ki-67  $< 14\%$ .

Luminal B/HER2 negative: ER positive and HER2 negative, and at least one of: Ki-67  $\geq 14\%$ , PR negative or low (cut-point  $< 20\%$ ).

Luminal B/HER2 positive: ER positive over-expressed or amplified, any PR, and any Ki-67.

HER2 positive: HER2 over-expressed or amplified, ER and PR absent.

TN (triple negative): ER and PR absent, and HER2 negative.

### ***WFO and concordance determination***

IBM WFO developed in collaboration with MSKCC, which is a cognitive computing system able to extract structured data from free text documents using natural language processing (NLP). It is a technology platform that uses NLP and machine learning to reveal insights from large amounts of unstructured data. It includes text from more than 300 medical journals and textbooks, MSKCC treatment guideline, and literature hand-selected by MSK experts.

Patients' data, including patient characteristics, tumor characteristics and stage, and some laboratory finding, were abstracted from electronic medical record system and entered manually into WFO by the trained senior oncology fellows.

After processing, chemotherapy recommendations were provided in 3 categories, green represents chemotherapy regimens "Recommended" with a strong base of evidence; amber represents chemotherapy regimens "For Consideration" which oncologists may consider as suitable alternatives base on their clinical judgment; and red represents chemotherapy regimens "Not Recommended", due to strong evidence against their use or some specific contraindications. Concordance was analyzed by comparing the decisions made by the oncologists to those proposed by WFO. Concordance was achieved when oncologists' treatment suggestions were in the "Recommended" or "For Consideration" categories given by WFO. WFO and the physicians who ran the cases were blinded to the chemotherapy regimens determined by oncologists.

### ***Data analysis and statistics***

The primary study objective was to evaluate the concordance between chemotherapy regimens oncologists' suggested and

WFO proposed. Descriptive statistics of breast cancer case characteristics were presented as means  $\pm$  standard deviation or median (range). Categorical variables were expressed as percentage (%). The concordance was expressed as percent agreement. Controlling for cancer characteristics, including patient age, menopausal status, tumor TNM stage, and molecular subtype, a logistic regression model of concordance between WFO and oncologists' suggestions was estimated with odds ratios (ORs) and 95% confidence intervals (CIs).

All statistical analyses were performed using Stata version 11.0 (StataCorp, College Station, TX, USA), and a significant difference was concluded for  $P < 0.05$ .

## Results

### Baseline characteristics

A total of 1,301 female patients diagnosed with stage I to III breast cancer between June 2013 and December 2017 were enrolled in this retrospective study. The clinicopathological characteristics of the total sample of breast cancer patients who received AC after surgery from June 2013 to December 2017 are reported in *Table 1*. The mean age at the time of chemotherapy started was 51.1 years. Of the 1,121 patients, pathological TNM stage and molecular subtype varied slightly, with more patients having stage II tumors (46.3%), and more patients having Luminal B (HER2 negative) tumors (52.1%).

*Table 2* shows the characteristics of 180 breast cancer patients who received NAC after biopsy from January 2014 to October 2017. The mean age at the time of NAC was 48.2 years. Similar to patients in the AC group, more patients presented with Luminal B (HER2 negative) tumors (39.4%) and stage II tumors (87.8%) in this group.

### Concordance analysis

Overall chemotherapy regimen concordance was 69.4%, with 65.0% in AC group and 96.7% in NAC group (*Table 3*). In subset analysis of breast cancer with respect to TNM stage and molecular subtype, the concordance varied greatly. AC recommendations were concordant in 92.3% of stage III breast cancer and 50.8% of stage I breast cancer. Moreover, the concordance varied by molecular subtype, which was higher for triple negative breast cancer (89.3%) than Luminal A and Luminal B/HER2 negative breast cancer (55.3% and 52.7%, respectively) (*Table 3, Figure 2*). Patients'

age also played a vital role in chemotherapy concordance rate. The chemotherapy regimen concordance declined significantly with increasing age, except for the age group 41–50 years (*Table 3, Figure 3*). As for chemotherapy regimen in NAC group, subset analyses of treatment concordance by stage and molecular subtype were also carried out. The regimen concordances were high in all the subgroups, which narrowly ranged from 93.1% to 100% (*Table 3, Figure 2*).

Results from logistic regression of concordance are presented in *Table 4*. ORs were analyzed by patient age, menopausal status, TNM stage, Molecular subtype, Ki-67 status, and vessel invasive status. In univariate analysis, old age and postmenopausal were associated significantly lower concordance rates. High TNM stage, high Ki-67 expression (Ki-67 >50%), and invasive vessel were demonstrated to increase the concordance (all  $P < 0.05$ ). ORs of concordance by molecular subtype, showed that compared with Luminal A cancers, Luminal B/HER2 positive, HER2 positive, and triple negative cancers were significantly more likely to be concordant (all  $P < 0.05$ ).

Adjusted for potential confounders in multivariate analysis, concordance was only significantly lower for patients 70 years older, compared with patients 40 years of age and younger (OR =0.33, 95% CI, 0.14–0.78,  $P = 0.012$ ). Stage II and III diseases were still significantly more likely to concordant than stage I disease (all  $P < 0.05$ ). Additionally, multivariate analysis revealed that only triple-negative breast cancers were significantly more likely to be concordant than Luminal A ( $P < 0.001$ ).

### Disagreements in AC population

*Table 5* presents the distribution of TNM stage and molecular subtype of non-concordant cases according to different AC regimen used in clinic. First, we analyzed the 245 patients who received chemotherapy without HER2 target therapy. Of the 245 patients, 41 patients received AC (doxorubicin and cyclophosphamide) chemotherapy, 106 patients received TC (docetaxel and cyclophosphamide), 30 patients received FEC (fluorouracil, epirubicin and cyclophosphamide), 36 patients received FEC-T (3 cycles of FEC followed by 3 cycles of docetaxel), and 32 patients received AC-T (4 cycles of AC followed by 4 cycles of paclitaxel) as oncologists suggested in clinic, while WFO recommended CMF (cyclophosphamide, methotrexate, and fluorouracil), ddAC-T (dose-dense AC-T), or no chemotherapy. It is obvious that most of the patients were

**Table 1** Characteristics of breast cancer cases receiving adjuvant chemotherapy after primary surgery

Variable	Number
Female, n (%)	1,121 [100]
Age, mean $\pm$ SD, years	51.1 $\pm$ 11.1
Menopausal status, n (%)	
Pre-menopausal	577 (51.47)
Post-menopausal	544 (48.53)
Tumor location, n (%)	
Left	572 (51.03)
Right	549 (48.97)
Surgery date (year), n (%)	
~2014	138 (12.31)
2015	368 (32.83)
2016	416 (37.11)
2017	199 (17.75)
Surgery type, n (%)	
Surgery for primary site	
Conserving surgery	247 (22.03)
Mastectomy	874 (77.97)
Without plastic surgery	833 (74.31)
With plastic surgery	41 (3.66)
Surgery for axillary lymph node	
SLNB	449 (40.05)
ALND	567 (50.58)
SLNB + ALND	105 (9.37)
Histopathology, n (%)	
IDC	1,060 (94.56)
ILC	30 (2.68)
Other	31 (2.77)
Pathological T stage, n (%)	
pT1	562 (50.13)
pT2	528 (47.10)
pT3	31 (2.77)
Pathological N stage, n (%)	
pN0	663 (59.14)
pN1	254 (22.66)
pN2	120 (10.70)
pN3	84 (7.49)

**Table 1** (continued)**Table 1** (continued)

Variable	Number
Pathological TNM stage, n (%)	
I	394 (35.15)
II	519 (46.30)
III	208 (18.55)
ER status, n (%)	
Positive	307 (27.39)
Negative	814 (72.61)
PR status, n (%)	
Positive	718 (64.05)
Negative	403 (35.95)
HER2 status, n (%)	
Positive	293 (26.14)
Negative	828 (73.86)
Ki-67, n (%)	
<14%	119 (10.62)
14–50%	617 (55.04)
>50%	385 (34.34)
Molecular subtype, n (%)	
Luminal A	85 (7.58)
Luminal B/HER2 negative	584 (52.10)
Luminal B/HER2 positive	162 (14.45)
HER2 positive	131 (11.69)
Triple negative	159 (14.18)
Vessel invasive status, n (%)	
Yes	218 (19.45)
No	903 (80.55)

SLNB, sentinel lymph node biopsy; ALND, axillary lymph node dissection; ILC, invasive lobular carcinoma; IDC, invasive ductal carcinoma; ER, estrogen receptor; PR, progesterone receptor.

Luminal A or Luminal B/HER2 negative irrespective of tumor stage, except for a small number of patients with HER2 positive or triple negative but stage I cancer.

For Luminal B/HER2 positive and HER2 positive breast cancers, 10 patients received AC-TH (AC-T plus trastuzumab) chemotherapy, 17 patients received FEC-TH (FEC-T plus trastuzumab), 26 patients received TCH (docetaxel and carboplatin plus trastuzumab), and 5 patients received PH (weekly paclitaxel and trastuzumab for

**Table 2** Characteristics of breast cancer cases received neoadjuvant chemotherapy after diagnosis

Variable	Number
Female, n (%)	180 [100]
Age, mean ± SD, years	48.2±11.0
Menopausal status	
Pre-menopausal	102 (56.67)
Post-menopausal	78 (43.33)
Tumor location	
Left	94 (52.22)
Right	86 (47.78)
Biopsy date (year), n (%)	
~2014	23 (12.78)
2015	43 (23.89)
2016	79 (43.89)
2017	35 (19.44)
Histopathology, n (%)	
Invasive breast cancer	180 [100]
Clinical T stage, n (%)	
cT1	21 (11.67)
cT2	139 (77.22)
cT3	20 (11.11)
Clinical N stage, n (%)	
cN0	67 (37.22)
cN1	106 (58.89)
cN2	7 (3.89)
Clinical TNM stage, n (%)	
I	–
II	158 (87.78)
III	22 (12.22)
Biopsy ER status, n (%)	
Positive	120 (66.67)
Negative	60 (33.33)
Biopsy PR status, n (%)	
Positive	96 (53.33)
Negative	84 (46.67)

**Table 2** (continued)**Table 2** (continued)

Variable	Number
Biopsy HER2 status, n (%)	
Positive	72 (40.00)
Negative	108 (60.00)
Biopsy Ki-67, n (%)	
<14%	9 (5.00)
14–50%	130 (72.22)
>50%	41 (22.78)
Biopsy molecular subtype, n (%)	
Luminal A	8 (4.44)
Luminal B/HER2 negative	71 (39.44)
Luminal B/HER2 positive	41 (22.78)
HER2 positive	31 (17.22)
Triple negative	29 (16.11)

ER, estrogen receptor; PR, progesterone receptor.

12 weeks), different from WFO recommended. Among these 58 patients, 31 (53.4%) patients were stage I, 26 (44.8%) patients were stage II, and only 1 (0.2%) patient was stage III.

It is worth to note that a total of 89 patients did not receive chemotherapy in clinic, and all of them were Luminal A and Luminal B1, except for one patient with HER2 positive who rejected chemotherapy treatment.

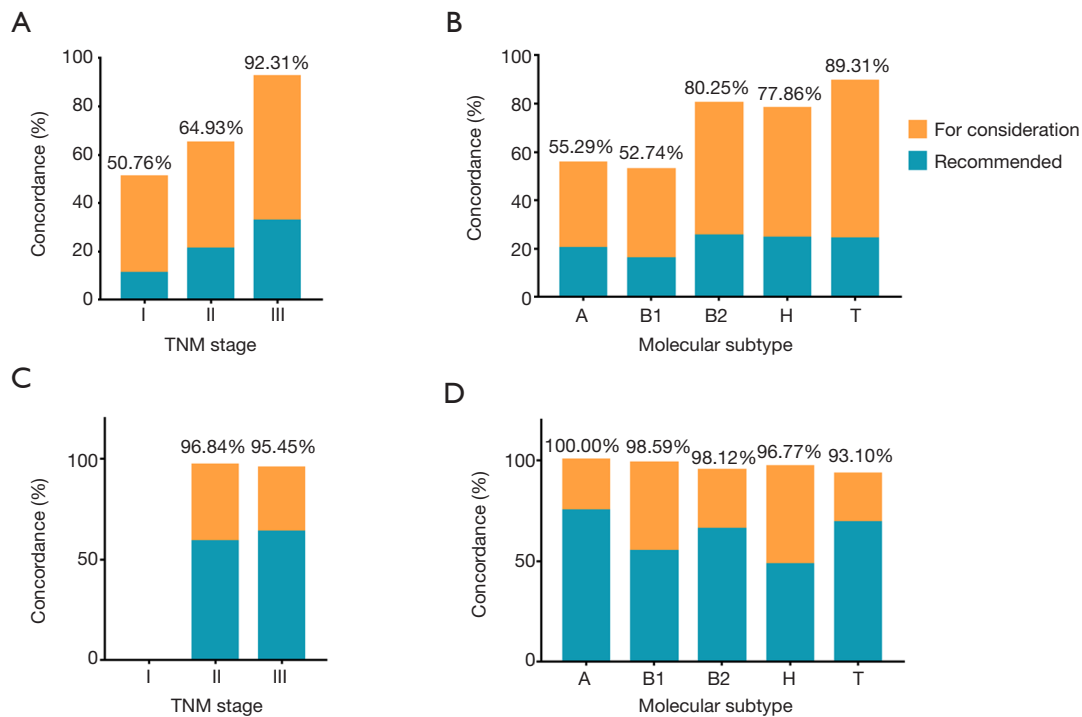
## Discussion

This large population based retrospective, observational study shows that chemotherapy regimen options suggested by WFO were concordant with the therapeutic decisions by oncologists (clinical use) in the large majority of breast cancer patients treated with NAC before surgery. The degrees of concordance in patients treated with AC was much lower, reflecting the differences in practice patterns between the United States, (where WFO was trained) and The First Affiliated Hospital with Nanjing Medical University in China. In subgroup analysis, there were still higher degrees of concordance between WFO treatments options and the oncologists' decision among stage III, HER2 positive, and triple negative breast cancers. However,

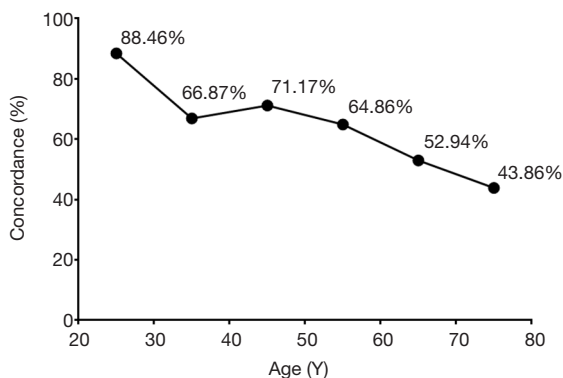
**Table 3** Clinical use and WFO recommendations

Review of breast cancer cases	Concordant cases, n (%)		
	Recommended	For consideration	Total
Total (n=1,301)	327 (25.13)	576 (44.27)	903 (69.41)
Adjuvant chemotherapy (n=1,121)	220 (19.63)	509 (45.41)	729 (65.03)
Surgery date, year			
~2014 (n=138)	35 (25.36)	51 (36.96)	86 (62.32)
2015 (n=368)	71 (19.29)	169 (45.92)	240 (65.22)
2016 (n=416)	79 (18.99)	197 (47.36)	276 (66.35)
2017 (n=199)	35 (17.59)	92 (46.23)	127 (63.82)
Age, years			
≤40 (n=189)	42 (22.22)	90 (47.62)	132 (69.84)
41–50 (n=392)	75 (19.13)	204 (52.04)	279 (71.17)
51–60 (n=313)	60 (19.17)	143 (45.69)	203 (64.86)
61–70 (n=170)	31 (18.24)	59 (34.71)	90 (52.94)
≥71 (n=57)	12 (21.05)	13 (22.81)	25 (43.86)
Pathologic TNM stage			
I (n=394)	43 (10.91)	157 (39.85)	200 (50.76)
II (n=519)	109 (21.00)	228 (43.93)	337 (64.93)
III (n=208)	68 (32.69)	124 (59.62)	192 (92.31)
Molecular subtype			
Luminal A (n=85)	17 (20.00)	30 (35.29)	47 (55.29)
Luminal B/HER2 negative (n=584)	92 (15.75)	216 (36.99)	308 (52.74)
Luminal B/HER2 positive (n=162)	41 (25.31)	89 (54.94)	130 (80.25)
HER2 positive (n=131)	32 (24.43)	70 (53.44)	102 (77.86)
Triple negative (n=159)	38 (23.90)	104 (65.41)	142 (89.31)
Neoadjuvant chemotherapy (n=180)	107 (59.44)	67 (37.22)	174 (96.67)
Biopsy date, year			
~2014 (n=23)	16 (69.57)	5 (21.74)	21 (91.30)
2015 (n=43)	34 (79.07)	9 (20.93)	43 (100.00)
2016 (n=79)	44 (55.70)	33 (41.77)	77 (97.47)
2017 (n=35)	13 (37.14)	20 (57.14)	33 (94.29)
Clinical TNM stage			
II (n=158)	93 (58.86)	60 (37.97)	153 (96.84)
III (n=22)	14 (63.64)	7 (31.82)	21 (95.45)
Molecular subtype			
Luminal A (n=8)	6 (75.00)	2 (25.00)	8 (100.00)
Luminal B/HER2 negative (n=71)	39 (54.93)	31 (43.66)	70 (98.59)
Luminal B/HER2 positive (n=41)	27 (65.85)	12 (29.27)	39 (95.12)
HER2 positive (n=31)	15 (48.39)	15 (48.39)	30 (96.77)
Triple negative (n=29)	20 (68.97)	7 (24.14)	27 (93.10)

WFO, Watson for Oncology.



**Figure 2** Chemotherapy regimen concordance between WFO and Clinical use by breast cancer stage and molecular subtype in AC group (A,B) and NAC group (C,D). WFO, Watson for Oncology; AC, adjuvant chemotherapy; NAC, neoadjuvant chemotherapy therapy; A, Luminal A; B1, Luminal B/HER2 negative; B2, Luminal B/HER2 positive; H, HER2 positive; T, triple-negative.



**Figure 3** Chemotherapy regimen concordance between WFO and clinical use by age in AC group. AC, adjuvant chemotherapy; WFO, Watson for Oncology.

with respect to Luminal A and Luminal B/HER2 negative disease, lots of improvements are needed from WFO.

We speculated that potential reasons for non-concordance included: (I) local medications or medication combinations not available in WFO; (II) differences in criteria for use of certain chemotherapy regimen; (III)

oncologist preference; (IV) patient and family preference; (V) Ki-67 is an important factor in determining molecular subtype, which is not considered in WFO. We will discuss the non-concordant regimen in details below.

In 1976, the efficacy of CMF was first reported, as adjuvant treatment for node positive breast cancer (17). Its long-term results represent an important step in the contemporary evolution of breast cancer treatment. And then CMF has been expanded to patients with node negative cancer (18). As recommended by WFO, a majority of cases were suggested to receive CMF chemotherapy. However, CMF has not been used for years in our hospital, or in most of cities in China.

Doxorubicin was first evaluated in NSABP B-11 trial (19). Later in 1990, NSABP B-15 trial was carried out in 2,194 patients with node-positive disease to AC every 3 weeks for 4 cycles versus CMF for 6 cycles, which demonstrated similar 3-year DFS rates and OS rates (20). Subsequently, NSABP B-23 also found no difference in outcomes in patients with node-negative disease who were treated with AC or CMF (21). Additionally, results from the US Oncology Research phase III trial demonstrated that



**Table 4** Logistic regression model of concordance between WFO and clinical use for patients receiving adjuvant chemotherapy after surgery (N=1,301)

Variable	Univariate Cox regression		Multivariable Cox regression	
	Odds ratio (95% CI)	P value	Adjusted odds ratio (95% CI)	P value
<b>Age at diagnosis, years</b>				
≤40 (reference)	1.00		1.00	
41–50	1.07 (0.73–1.56)	0.741	1.10 (0.72–1.69)	0.665
51–60	0.80 (0.54–1.17)	0.251	0.99 (0.54–1.81)	0.964
61–70	0.49 (0.32–0.75)	0.001	0.61 (0.31–1.21)	0.159
70+	0.34 (0.18–0.629)	<0.001	0.33 (0.14–0.78)	0.012
<b>Menopausal status</b>				
Premenopausal (reference)	1.00		1.00	
Postmenopausal	0.61 (0.47–0.78)	<0.001	0.69 (0.41–1.13)	0.140
<b>Pathological TNM stage</b>				
I (reference)	1.00		1.00	
II	1.80 (1.37–2.35)	<0.001	2.12 (1.57–2.88)	<0.001
III	11.64 (6.74–20.11)	<0.001	13.87 (7.66–25.11)	<0.001
<b>Ki-67</b>				
<14% (reference)	1.00		1.00	
14–50%	1.28 (0.86–1.90)	0.222	1.44 (0.69–3.02)	0.329
>50%	2.86 (1.86–4.40)	<0.001	1.75 (0.81–3.75)	0.154
<b>Molecular subtype</b>				
Luminal A (reference)	1.00		1.00	
Luminal B/HER2 negative	0.90 (0.57–1.43)	0.659	0.54 (0.23–1.26)	0.153
Luminal B/HER2 positive	3.28 (1.85–5.85)	<0.001	1.77 (0.70–4.51)	0.23
HER2 positive	2.84 (1.57–5.15)	0.001	1.80 (0.70–4.60)	0.223
Triple negative	6.75 (3.49–13.07)	<0.001	5.95 (2.25–15.75)	<0.001
<b>Vessel invasive status</b>				
No (reference)	1.00		1.00	
Yes	2.32 (1.63–3.30)	<0.001	1.44 (0.954–2.17)	0.082

Potential confounders: age, menopausal status, TNM stage, molecular subtype, Ki-67 status, and vessel invasive status. WFO, Watson for Oncology.

TC was associated with significantly improved outcome compared with AC (22). As *Table 5* presents, when WFO recommended CMF, oncologists chose AC or TC for some cases in our study. AC and TC are two common chemotherapy regimens in our clinic practice when the disease is in early stage or is a Luminal type in low risk.

The two trials, FASG 2 and 7 found that there was a

trend in favor of FEC for OS compared with tamoxifen alone (23). The EBCTCG meta-analysis in 2012 found that FEC was more effective in reducing breast cancer mortality compared to CMF (RR =0.78, P=0.0004) (2). In addition, the PACS01 trial evaluated FEC with FEC-T, which found that FEC-T was associated with improved DFS (HR =0.85, P=0.036) and OS (HR =0.75, P=0.007) (24). FEC and

**Table 5** Characteristics of non-concordant cases according to different chemotherapy regimens used in clinic

Clinical chemotherapy regimen	WFO chemotherapy regimen	Molecular subtype	TNM stage		
			I	II	III
AC [41]	CMF/ddAC-T/none	A	3	3	–
		B1	23	10	1
		H	1 (0.2 cm)	–	–
AC-T [32]	CMF/ddAC-T/none	B1	17	9	–
		T	6	–	–
TC [106]	CMF/ddAC-T/none	A	1	4	1
		B1	27	60	4
		H	1 (0.1 cm)	–	–
		T	7 (0.4–1.8 cm)	1 (76 y)	–
FEC [30]	CMF/ddAC-T	B1	14	16	–
FEC-T [36]	CMF/ddAC-T	B1	11	22	–
		T	2	1	–
AC-TH [10]	TH [1]/none	B2	3 (0.2–0.7 cm)	–	–
		H	7 (0.1–0.5 cm)	–	–
FEC-TH [17]	TH/ddAC-TH/none	B2	5 (0.5–1.8 cm)	7	–
		H	2 (0.2–0.3 cm)	3	–
TCH [26]	ddAC-TH/TH	B2	6 (0.2–1.8 cm)	10	–
		H	4	5	1
PH [5]	ddAC-TH/none	B2	1	–	–
		H	3 (0.1–0.3 cm)	1 (70 y)	–
None [89]	CMF/ddAC-T	A	17	9	–
		B1	33	23	6 (60+ y)
		H	1	–	–

WFO, Watson for Oncology; AC, doxorubicin and cyclophosphamide; AC-T, 4 cycles of AC followed by 4 cycles of paclitaxel; TC, docetaxel and cyclophosphamide; FEC, fluorouracil, epirubicin and cyclophosphamide; FEC-T, 3 cycles of FEC followed by 3 cycles of docetaxel; AC-TH, AC-T plus trastuzumab; FEC-TH, FEC-T plus trastuzumab; TCH, docetaxel and carboplatin plus trastuzumab; PH, weekly paclitaxel and trastuzumab for 12 weeks; CMF, cyclophosphamide, methotrexate, and fluorouracil; ddAC-T, dose-dense AC-T.

FEC-T were two common regimens used before the year 2017 in our hospital, but never recommended by WFO.

CALGB 9344 (25) and NSABP B-28 (26) both demonstrated AC-T was associated with better outcome than AC. Afterwards, results of C9741 after a median 6.5 years of follow-up favored ddAC-T in DFS and OS in ER-negative disease but not ER-positive disease (27). AC-T and ddAC-T are two favorite regimens in WFO, which we always prefer for patients in high risk, including patients

with young age, large tumor size, positive lymph node, and Triple-negative subtype.

If HER2 is positive, trastuzumab will be added to the regimens, like FEC-TH, AC-TH and ddAC-TH. These are common-use regimens in clinical practice. The results of BCIRG-006 trial showed that risk-benefit ration favored the non-anthracycline TCH regimen over AC-TH, given its similar efficacy and fewer toxic effects (28). So, for some cases that WFO recommend ddAC-TH, the oncologists

preferred TCH instead, which is not included in WFO. Additionally, PH is another choice for oncologists in clinical practice. An uncontrolled, single-group study in 406 patients resulted that patients with predominantly stage I, HER2 positive, and node negative breast cancer, treatment with PH correlated with a risk of early recurrence of about 2% (29). According to the evidence above, oncologists usually chose the PH regimen for stage I and HER2 positive patients.

Notably, there were 89 patients who did not receive any chemotherapy. Nearly all the patients were Luminal A or Luminal B1. These are endocrine responsive breast cancers, and most of them would get no significant beneficial effect from chemotherapy.

Several limitations of this study should be considered. First, the present study was a retrospective, observational study without any controls. And we cannot compare the efficacy of regimens suggested by oncologists and WFO recommendations. Therefore, lack of concordance does not necessarily provide evidence regarding whether the clinical regimen or WFO was “correct” in its recommendation. Second, it is better to assess the Oncotype DX Recurrence Score to identify which group of the patients are in high risk and would benefit from chemotherapy, especially for Luminal A and Luminal B/HER2 negative disease. No multigene assay was used in making clinical decisions in our study. Third, reasons for non-concordance such as insurance requirements, cost, and patient and oncologist preferences were not analyzed in our study.

## Conclusions

WFO is a step towards personalized medicine. It could be an essential tool to oncologists by reducing the cognitive burden of physicians in keeping up with medical literature by providing clinically actionable insights to assist them in treating patients. Chemotherapy regimens provided by WFO did not exhibit a high degree of agreement with those suggested by oncologists in clinical practice in the hospital in China. We supposed that WFO's capabilities as a cognitive decision support can be further improved by an incorporating regional guideline, enabling oncologists and patients to benefit from WFO worldwide. However, it should be kept in mind that WFO will be only an assisting tool and it will never be able to replace the patient-doctor relationship, which is a very essential component in treating patients suffering with cancer.

## Acknowledgments

*Funding:* This work was supported by the National Natural Science Foundation of China [81502286, 81502299, 81572607 and 81771953], the 2016 Industry Prospecting and Common Key Technology Key Projects of Jiangsu Province Science and Technology Department [BE2016002-4], the 2016 projects of Nanjing Science Bureau [201608003], and a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

## Footnote

*Conflicts of Interest:* Both authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/tcr.2019.01.34>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study protocol was approved by the ethics committee of The First Affiliated Hospital with Nanjing Medical University (2018-SR-368) and was conducted in accordance with the Helsinki Declaration of 1964 (as revised in 2008). Due to the retrospective nature of this study, the need for informed consent was waived.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. *CA Cancer J Clin* 2015;65:87-108.
2. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Peto R, Davies C, et al. Comparisons between different polychemotherapy regimens for early

- breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. *Lancet* 2012;379:432-44.
3. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* 2015;386:1341-52.
  4. Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005;353:1673-84.
  5. Joensuu H, Kellokumpu-Lehtinen PL, Bono P, et al. Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. *N Engl J Med* 2006;354:809-20.
  6. Fisher B, Ravdin RG, Ausman RK, et al. Surgical adjuvant chemotherapy in cancer of the breast: results of a decade of cooperative investigation. *Ann Surg* 1968;168:337-56.
  7. Fisher B, Anderson S, Bryant J, et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *N Engl J Med* 2002;347:1233-41.
  8. Munoz D, Near AM, van Ravesteyn NT, et al. Effects of screening and systemic adjuvant therapy on ER-specific US breast cancer mortality. *J Natl Cancer Inst* 2014. doi: 10.1093/jnci/dju289.
  9. Albain KS, Barlow WE, Shak S, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. *Lancet Oncol* 2010;11:55-65.
  10. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 2006;24:3726-34.
  11. Hammond ME, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol* 2010;28:2784-95.
  12. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. *N Engl J Med* 2009;360:790-800.
  13. Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* 2013;31:3997-4013.
  14. Sparano JA, Fazzari M, Kenny PA. Clinical application of gene expression profiling in breast cancer. *Surg Oncol Clin N Am* 2010;19:581-606.
  15. Woolhandler S, Himmelstein DU. Administrative work consumes one-sixth of U.S. physicians' working hours and lowers their career satisfaction. *Int J Health Serv* 2014;44:635-42.
  16. Somashekhar SP, Sepulveda MJ, Puglielli S, et al. Watson for Oncology and breast cancer treatment recommendations: agreement with an expert multidisciplinary tumor board. *Ann Oncol* 2018;29:418-23.
  17. Bonadonna G, Brusamolino E, Valagussa P, et al. Combination chemotherapy as an adjuvant treatment in operable breast cancer. *N Engl J Med* 1976;294:405-10.
  18. Mansour EG, Gray R, Shatila AH, et al. Survival advantage of adjuvant chemotherapy in high-risk node-negative breast cancer: ten-year analysis--an intergroup study. *J Clin Oncol* 1998;16:3486-92.
  19. Fisher B, Redmond C, Wickerham DL, et al. Doxorubicin-containing regimens for the treatment of stage II breast cancer: The National Surgical Adjuvant Breast and Bowel Project experience. *J Clin Oncol* 1989;7:572-82.
  20. Fisher B, Brown AM, Dimitrov NV, et al. Two months of doxorubicin-cyclophosphamide with and without interval reinduction therapy compared with 6 months of cyclophosphamide, methotrexate, and fluorouracil in positive-node breast cancer patients with tamoxifen-nonresponsive tumors: results from the National Surgical Adjuvant Breast and Bowel Project B-15. *J Clin Oncol* 1990;8:1483-96.
  21. Fisher B, Anderson S, Tan-Chiu E, et al. Tamoxifen and chemotherapy for axillary node-negative, estrogen receptor-negative breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-23. *J Clin Oncol* 2001;19:931-42.
  22. Jones S, Holmes FA, O'Shaughnessy J, et al. Docetaxel With Cyclophosphamide Is Associated With an Overall Survival Benefit Compared With Doxorubicin and Cyclophosphamide: 7-Year Follow-Up of US Oncology Research Trial 9735. *J Clin Oncol* 2009;27:1177-83.
  23. Namer M, Fargeot P, Roche H, et al. Improved disease-free survival with epirubicin-based chemoendocrine adjuvant therapy compared with tamoxifen alone in one to three node-positive, estrogen-receptor-positive, postmenopausal breast cancer patients: results of French Adjuvant Study Group 02 and 07 trials. *Ann Oncol*

- 2006;17:65-73.
24. Coudert B, Asselain B, Campone M, et al. Extended benefit from sequential administration of docetaxel after standard fluorouracil, epirubicin, and cyclophosphamide regimen for node-positive breast cancer: the 8-year follow-up results of the UNICANCER-PACS01 trial. *Oncologist* 2012;17:900-9.
  25. Henderson IC, Berry DA, Demetri GD, et al. Improved outcomes from adding sequential Paclitaxel but not from escalating Doxorubicin dose in an adjuvant chemotherapy regimen for patients with node-positive primary breast cancer. *J Clin Oncol* 2003;21:976-83.
  26. Mamounas EP, Bryant J, Lembersky B, et al. Paclitaxel after doxorubicin plus cyclophosphamide as adjuvant chemotherapy for node-positive breast cancer: results from NSABP B-28. *J Clin Oncol* 2005;23:3686-96.
  27. Citron ML, Berry DA, Cirincione C, et al. Randomized trial of dose-dense versus conventionally scheduled and sequential versus concurrent combination chemotherapy as postoperative adjuvant treatment of node-positive primary breast cancer: first report of Intergroup Trial C9741/ Cancer and Leukemia Group B Trial 9741. *J Clin Oncol* 2003;21:1431-9.
  28. Slamon D, Eiermann W, Robert N, et al. Adjuvant trastuzumab in HER2-positive breast cancer. *N Engl J Med* 2011;365:1273-83.
  29. Tolaney SM, Barry WT, Dang CT, et al. Adjuvant paclitaxel and trastuzumab for node-negative, HER2-positive breast cancer. *N Engl J Med* 2015;372:134-41.

**Cite this article as:** Pan H, Tao J, Qian M, Zhou W, Qian Y, Xie H, Jing S, Xu T, Zhang X, Dai Z, You M, Liu Y, Liu X, Wang S. Concordance assessment of Watson for Oncology in breast cancer chemotherapy: first China experience. *Transl Cancer Res* 2019;8(2):389-401. doi: 10.21037/tcr.2019.01.34