

# 日本語版 Quality Assessment Tool for Quantitative Studies (J- QATQS) の作成および信頼性の検討－等価性の担保を中心として－

細 坂 泰 子<sup>1</sup>      柏 崎 真 由<sup>2</sup>

<sup>1</sup> 東京慈恵会医科大学医学部看護学科

<sup>2</sup> 東京慈恵会医科大学大学院医学研究科看護学専攻修士課程

(受付 2018 年 8 月 13 日)

## CREATION OF A JAPANESE-LANGUAGE VERSION OF THE QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES (J-QATQS) ENSURING EQUIVALENCE, AND A STUDY OF ITS RELIABILITY

Yasuko Hososaka<sup>1</sup>, Mayu Kashiwazaki<sup>2</sup>

<sup>1</sup>*The Jikei University School of Nursing*

<sup>2</sup>*Master's Program in Nursing, The Jikei University Graduate School of Nursing*

In the present study, the Quality Assessment Tool for Quantitative Studies (QATQS), a scale for comprehensively evaluating the quality of quantitative research, was translated to create a Japanese-language version (J-QATQS); the equivalence of the two versions was ensured, and the reliability of the Japanese version was examined. The translation of the QATQS was evaluated with the back-translation method. The J-QATQS was created in a chart format to ensure ease of use. Additionally, in the constructive concept of the QATQS, only widely recognized statistical terms are used, and because we believe no problems of validity were present, only inter-rater reliability was analyzed. Of the nursing studies written in Japanese and published by two researchers on the Ichushi-Web, a Internet version of the Japan Medical Abstracts Society, 10 studies of comparative testing and comparative research were evaluated with the J-QATQS, and the inter-rater reliability ( $\kappa$  coefficient) was calculated. In all cases, the  $\kappa$  coefficient satisfied the level of significance, and reliability was found to be high for all of the studies on the basis of the following scores: selection bias,  $\kappa=0.94$ ; research design,  $\kappa=0.86$ ; confounding factor,  $\kappa=1$ ; blinded,  $\kappa=0.67$ ; data collection method,  $\kappa=0.68$ ; and target dropout rate,  $\kappa=1$ . The equivalence of the J-QATQS created in the present study was confirmed by the authors of the QATQS, and the results of the reliability study showed that the inter-rater reliability of the J-QATQS was high, indicating the possibility that published studies can be objectively evaluated.

(Tokyo Jikeikai Medical Journal 2019;134:75-81)

Key words : Japanese-language version of Quality Assessment Tool for Quantitative Studies, quantitative studies, reviews

## I. 緒 言

研究にエビデンスは必要不可欠となっている。現在では国内外で医学研究結果を検索できるデータベースを活用することで、多くの知見に接することが可能になった。一方でそれらの研究データは玉石混交であり、有益なデータもあるものの研究デザインや研究手法に疑念を生じるものも少なくない。多くの文献を手当たり次第に読むことで情報は集められるが、それをどのように評価するのかについては母集団やバイアスの存在を個々の研究者が比較してレビューするしかなかった。

Quality Assessment Tool for Quantitative Studies (以下、QATQSと略す)は1998年に公衆衛生活動の実践をサポートするエビデンスとして質の高いsystematic reviewを提供するために作成された<sup>1)</sup>。選択バイアス、研究デザイン、交絡因子、盲検化、データ収集方法、対象の脱落率、介入方法、分析方法の8項目を検証し(評価に使用するのは介入方法、分析方法を除いた6項目のみ)最終的な判定を行うもので、研究者の評価の手助けとしてユーザーマニュアルも添付されている。このツールは反復プロセスを使用した有効性や検査の一貫性、構成概念妥当性が示されており、また2名の評価者の一致率が $\kappa$ 係数0.74および0.61と高い<sup>2)</sup>ことが明らかになっている。現在までにこのツールを用いたレビューが多く行われており<sup>3-5)</sup>、レビューを行う上での有効なツールであるといえる。

欧米では量的研究のレビュー評価を行うための尺度が作成されているが、本邦では日本語で利用できる一般的な尺度が存在しないため、本研究で日本語版QATQS(以下、J-QATQSとする)を作成することには意義がある。量的研究の客観的な評価を測定するためのツールとしてJ-QATQSを作成することによって、我が国の研究レビューのレベルを上げること、さらには評価ツールを用いた有意義なレビュー研究の推進を図ることが可能である。

本研究はQATQSとの等価性を担保したJ-QATQSの作成、およびその信頼性を検討することを目的とした。なお本ツールは通常検証が必要である構成概念としては、既に統計用語として

広く認知されている用語のみが使用されており問題は無いと考えられるため、評価者がJ-QATQSを用いて得られた評価者信頼性を検証した。

## II. 研究 方 法

### 1. 翻訳の対象尺度

今回使用したQATQSは選択バイアス、研究デザイン、交絡因子、盲検化、データ収集方法、対象の脱落率の6項目をそれぞれSTRONG, MODERATE, WEAKの3つに分けて判定し、論文のグローバルな格付けを行うことができるツールである。それぞれの項目ごとに1～2問の問いがあり、その選択によって評価が与えられる。評価は2名の査読者で行い、両者の判定を加味しながら最終的な格付けを行うことができるツールであり、1998年にEffective Public Health Practice Projectによって開発された。このQATQSおよびツールに付随するユーザーマニュアルを翻訳の対象尺度とした。日本語版作成にあたり、2017年8月にQATQSの原著者であるDr. Donna Fitzpatrick-Lewi, Drs. Donna Ciliska, Drs. Helen Thomasに日本語翻訳の許可を得た。

### 2. 翻訳手法の選択

海外の尺度翻訳については、デュアルパネル法(以下、DP)やバクトランスレーション法(以下、BT)が使用されることが多い。DPは翻訳の言語的一致よりも概念上の一致を重視することで文化的背景に配慮した表現をとることが可能な手法である<sup>6)</sup>。一方BTは翻訳の質を高め妥当性が高いとされている<sup>7)</sup>。本尺度はレビューの評価尺度であり、文化的背景や概念を対象としないため、BT法を採用した。

### 3. 翻訳の実施方法

BTは稲田のガイドライン<sup>8)</sup>を参考に、以下の手順で実施した。

#### 1) 順翻訳

英語から日本語への翻訳を行った。著者2名が独立して順翻訳を行い、原版の意味を損なうことなく、回答者が容易に意味を理解できるよう表現を使うことに配慮した。QATQSとユーザーマニュアルは独立して順翻訳を行ったが、その後、お互いの順翻訳版を討議・統合し、J-QATQSを作成

する際に、ユーザーマニュアルをツールの中に入れ込む形で一つの版を作成した。なお最終的な評価に寄与しない設問は削除した。

## 2) 逆翻訳と原著者の校正

順翻訳版の逆翻訳を元の尺度を見ることなくブラインドで行うために、英語およびレビュー研究に精通している業者（株式会社フォルテ）に依頼して、原版の英語に翻訳した。依頼の際はこの翻訳が尺度翻訳の逆翻訳であることを伝えたうえで依頼した。逆翻訳を行ったJ-QATQSは原版の著者に比較してもらった。修正が必要な部分に関しては順翻訳者と検討した。その後2カ国語の異なる言語間で生じる項目表現が等価であるかについて、原版の著者らに確認してもらうことで、J-QATQSの等価性について検討した。また項目の概念に異なる乖離が生じていないかについても確認してもらった。

## 3) 認知デブリーフィングと校正

修正したJ-QATQSについて、文言の理解しやすさ、認知的等価性を検討するために日本語を母語とし、尺度を用いた研究実践の経験のある教員3名および院生2名を研究協力者として認知デブリーフィングを行った。これらの結果は総合し、用語の意味が分かりにくい箇所には用語の補足を入れるなど必要に応じて説明を付記した。また追記・修正部分については再度逆翻訳を行い、原著者の確認を行ったうえで、最終校正を行った。

## 4. 信頼性の検討

研究者2名が医学中央雑誌web版に掲載された和文の看護・医学研究のうち、比較試験・比較研究21編<sup>9-29)</sup>をそれぞれが独立してJ-QATQSで評価し、検者間信頼性（1次の重み付け $\kappa$ 係数）を算出した。STRONG, MODERATE, WEAKの順に重み付けを行い、交絡因子および盲検化の項目で「該当せず」に評価されたものはこの項目のみ結果をブランクとした。文献は「母性」と「周産期」のキーワードで2000年以降を検索年とし、絞り込み条件で原著論文、日本語、研究デザインとしてランダム化比較試験、準ランダム化比較試験、比較試験にチェックをして検索した。また学会誌に掲載されたものに限定した。最初にQATQSを使用した文献を見つけたのは母乳育児に関連する要因の文献検討<sup>3)</sup>であり、また、研究

者らは母性看護学領域の研究を行っているため、母性に関連する医学・看護学の文献を中心とした。研究デザインによる差異を考慮するため、コホート研究、ケースコントロール、分時型時系列など、できる限り異なる論文21編を抽出した。統計解析ソフトはSPSS for Windows Ver. 24.0を用い、有意水準は5%とした。

なお本研究で使用した論文は著作権法に則り、論文を収集し使用した。また研究にあたり本学倫理委員会に申請（31-070（9569））したが、人を対象としない研究であり、申請を要しないと判定された。

## III. 結 果

### 1) QATQSとの等価性を担保したJ-QATQSの作成

既存のQATQSは選択バイアス、研究デザイン、交絡因子、盲検化、データ収集方法、対象の脱落率の6項目を2～4つの質問とQATQSとは別のユーザーマニュアルから検証しなければならず、説明文書が長いことから使いにくい問題があった。また項目によっては結果がSTRONGともMODERATEとも取れる揺らぎのある割り付けがある場合があった。そのためJ-QATQSは使用者の利便性を鑑み、ユーザーマニュアルをJ-QATQSの中に入れ込み、項目ごとに2～3の質問で構成され、STRONG, MODERATE, WEAKのいずれかに集約される形のステップ・チャートを作成した（Fig. 1）。最終評価はWEAKの数が2項目以上の場合には、WEAK, 1項目ではMODERATE, ない場合はSTRONGとなるように作成した（Table 1）。作成したステップ・チャート形式のJ-QATQSについて、QATQS原著者のDr. Fitzpatrick-Lewis, Donnaに確認していただき、QATQSとの等価性について問題がないことを確認した。

### 2) J-QATQSの検者間信頼性

研究者2名で対象論文21編をそれぞれが評価した。Aの選択バイアスは2者の一致率が95%で

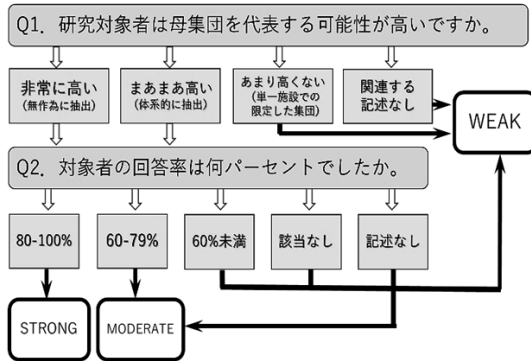
Table 1. Grobal Rating for This Paper

1	STRONG	WEAKの評価がない
2	MODERATE	WEAKが1つある
3	WEAK	2つ以上のWEAKがある

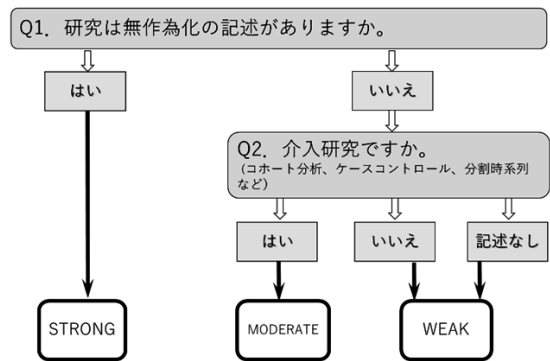
$\kappa$  係数は0.94であった。Bの研究デザインは一致率が95%で $\kappa$  係数は0.86であった。Cの交絡因子は検者間に差異がなく $\kappa$  係数は1であった。Dの盲検化は一致率が86%で $\kappa$  係数は0.67であった。

Eのデータ収集方法は一致率が71%で $\kappa$  係数は0.68であった。Fの対象の脱落率は一致率が71%で $\kappa$  係数は0.68であった。Fの対象の脱落率および最終評価は差異がなく、 $\kappa$  係数は1であった。p値はいずれも有意水準を満たした (Table 2)。

### A. 選択バイアス

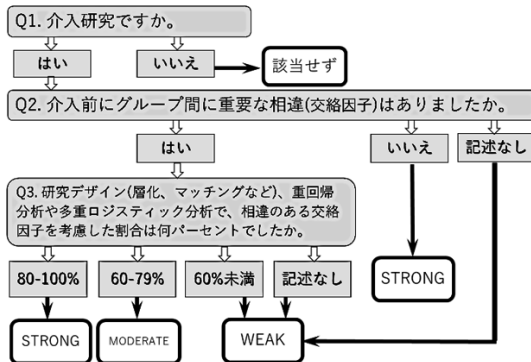


### B. 研究デザイン

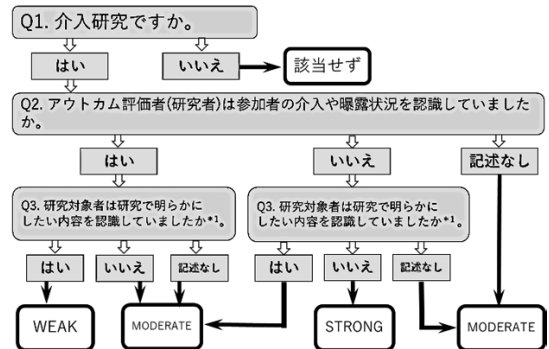


### C. 交絡因子

：人種、性別、婚姻状況/家族、年齢、収入、社会階級、教育、健康状態、アウトカム指標の介入前の値等

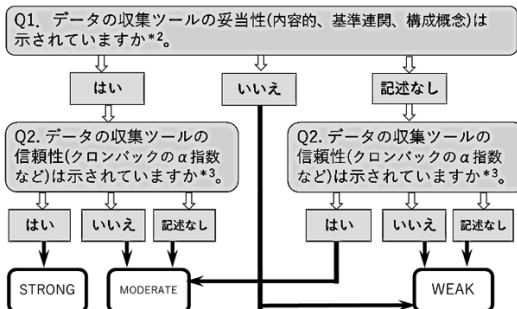


### D. 盲検化



\*1 後方視調査の場合、Q3は「いいえ (認識していない)」と判断する

### E. データ収集方法



\*2, 3 数値 (検査結果、モニター値など) で評価できる場合は「はい (信頼性・妥当性を示している)」と判断する

### F. 対象の脱落率

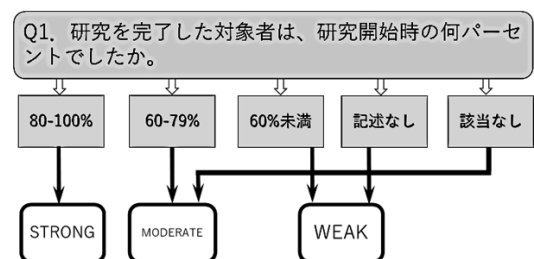


Fig. 1. The Japanese version (J-QATQS) chart format

Table 2. Inter-rater reliability of J-QATQS

項目	評価者A／評価者Bの評価数				2者間の一致率	$\kappa$ 係数	95%信頼区間	p
	STRONG	MODERATE	WEAK	該当せず				
A 選択バイアス	2/2	1/2	18/17	－	95%	0.94	0.86-0.98	p<0.001
B 研究デザイン	6/5	7/7	8/9	－	95%	0.86	0.69-0.94	p<0.001
C 交絡因子	3/3	0/0	8/8	10/10	100%	1	1	p<0.001
D 盲検化	2/3	6/3	3/5	10/10	86%	0.67	0.35-0.99	p<0.01
E データ収集方法	5/6	6/4	10/11	－	71%	0.68	0.44-0.92	p<0.001
F 対象の脱落率	16/16	4/4	1/1	－	100%	1	1	p<0.001
最終評価	0/0	2/2	18/18	－	100%	1	1	p<0.001

#### IV. 考 察

本研究で作成したJ-QATQSは原版のQATQSで使用する手引書をツールの中に入れ、評価の揺らぎをなくしたことで利便性の高いツールとなったと考えられる。このツールはステップ・チャート形式を用いた。ステップ・チャートは順序立てる、構造化する、要約する利点がある<sup>30)</sup>。分岐を持つステップを入れることで、複雑な流れをイメージでき全体の構造が理解しやすくなると考えられる。QATQSの作成者にもこのステップ・チャート形式での作成は等価性を保ちつつ、使用者の利便性を向上させたことについて評価された。

検者間信頼性についてはいずれの項目でも高い信頼性が得られた。盲検化やデータ収集方法でやや信頼性が低く、盲検化では倫理的配慮の記載が不十分であることから研究目的を対象に伝えているか不明瞭であることが、データ収集方法ではデータの信頼性・妥当性に関する記述が不十分であることから検者によって結果を想起する形での評価となる場合があること、研究方法欄に記載が必要な交絡因子や盲検化が結果に記載があることなどで検者が読み落とす場合があり、結果への判断が揺れることが原因と考えられた。

また本邦における看護研究はランダム化比較試験や準ランダム化比較試験、および比較試験の割合が3.4%と低い(医学中央雑誌でキーワードを看護研究とし、原著論文で検索された件数は37,835件であったが、そのうちのランダム化比較試験や準ランダム化比較試験、および比較試験の件数は1,298件：平成30年7月27日検索)。系統立てられた量的研究がやや乏しいこ

とから、J-QATQSでの最終評価が多くの場合でWEAKに分類される限界が示唆された。QATQSは公衆衛生看護領域で作成された尺度であることから、J-QATQSの検証論文を看護学を中心とした看護および医学研究に限定したが、上記の理由から和文での看護研究のみを対象にするには限界がある。しかし一部の看護研究や介入研究が多い医学研究や英文での評価には有効であると考えられ、J-QATQSの有用性が示唆された。

今後の課題として、より多くの論文でのJ-QATQSによる検者間信頼性を蓄積する必要がある。また今回の検者間信頼性を検討した研究者は量的研究の実践者であり、高等教育で看護研究の教育を受けていた。QATQSの使用については、原著者は研究における批判的吟味の専門技能を有する人を対象としている。しかし看護実践を行う場合のエビデンスを提供するためのスクリーニングに用いる場合、研究用語に熟知していない実践者が使用する可能性がある。ツールの用語はなるべく理解しやすいように説明文書をステップ・チャートの中に入れて作成したが、不十分な可能性がある。今後は理解しやすいツールの開発の検討を行う必要がある。

#### V. 結 語

本研究で作成したJ-QATQSは原著者による等価性を確認し、その信頼性を検討した結果、ツールの検者間信頼性は高く、論文を客観的に評価できる可能性がある。



# 著者の利益相反 (conflict of interest : COI) 開示 :

本論文の研究内容に関連して特に申告なし

## 文 献

- Effective Public Health Practice Project [Internet]. [https://merst.ca/wp-content/uploads/2018/02/quality-assessment-tool\\_2010.pdf](https://merst.ca/wp-content/uploads/2018/02/quality-assessment-tool_2010.pdf). [accessed 2018-07-27]
- Thomas BH, Ciliska D, Dobbins M, Micucci S. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evid Based Nurs*. 2004;1:176-84.
- Boccolini CS, Carvalho ML, Oliveira MI. Factors associated with exclusive breastfeeding in the first six months of life in Brazil: a systematic review. *Rev Saude Publica*. 2015; 49: 91.
- Dietz P, Watson ED, Sattler MC, Ruf W, Titze S, van Poppel M. The influence of physical activity during pregnancy on maternal, fetal or infant heart rate variability: a systematic review. *BMC Pregnancy Childbirth*. 2016; 16: 326.
- McPherson P, Krotofil J, Killaspy H. Mental health supported accommodation services: a systematic review of mental health and psychosocial outcomes. *BMC Psychiatry*. 2018; 18: 128.
- 宮本礼子, ボンジェ・ペイター, 須山夏加, 伊藤祐子. 保健医療系学生を対象とした「障害者に対する態度尺度日本版」の作成: ラッシュモデルを用いた信頼性・妥当性の検討. 作業療法. 2015;34:630-41.
- 福田由紀. 心理学実験における材料翻訳手順の妥当性. 法政大学文学部紀要. 2010;61:137-51.
- 稲田尚子. 尺度翻訳に関する基本指針. 行動療法研究. 2015;41:117-25.
- 前原邦江, 森恵美, 土屋雅子, 坂上明子, 岩田裕子, 小澤治美 ほか. 出産施設を退院後から産後1カ月までに母親役割の自信が高まる要因: 高年初産婦と34歳以下初産婦を比較して. 母性衛生. 2015;56:264-72.
- 田中静枝, 池内佳子. Hands-offテクニックによる母乳育児支援の効果. 母性衛生. 2013;54:275-85.
- 土取洋子. 乳児の授乳方法と離乳期の栄養に関する調査研究: 3歳児の母親を対象として. 母性衛生. 2005;45:445-53.
- 横山美江, 村井ちか子, 宮下茜, 辰巳朋美, 藤岡弘季. 授乳期の栄養方法の現状と母親の育児への思いに関する分析. 日本公衛誌. 2012;59:771-80.
- 島田三恵子, 杉本充弘, 縣俊彦, 新田紀枝, 関和男, 大橋一友 ほか. 産後1か月間の母子の心配事と子育て支援のニーズおよび育児環境に関する全国調査: 「健やか親子21」5年後の初経産別, 職業の有無による比較検討. 小児保健研究. 2006;65:752-62.
- 村上明美, 喜多里己, 神谷桂. 産褥早期の母親に対する癒しケアが産後の疲労と母乳育児に及ぼす影響. 日助会誌. 2008;22:136-45.
- 光武玲子, 新小田春美, 長谷川泰代, 荒田弘樹, 上野ふじ美, 諸隅誠一. NICU・GCUの低出生体重児における音楽介入の生理的影響とアクチグラフによる身体活動量・睡眠の質についての検討. 日新生児看会誌. 2017;23:16-22.
- 植松裕子, 祖父江育子. 新生児の静脈採血における24%シヨ糖液とおしゃぶりの鎮静効果: 一部二重盲検法無作為化比較試験. 日新生児看会誌. 2014;20:9-18.
- 大平肇子, 町浦美智子, 斉藤真, 村本敦子. 月経前症候群の症状を有する女性に対する呼吸法のリラクセーション効果. 母性衛生. 2013;53:497-504.
- 前田耕助, 大黒理恵, 大河原知嘉子, 樺島稔, 齋藤やよい. 電法による背部への温度刺激が脳血流動態に及ぼす影響. 日本看護技術学会誌. 2017;15:245-54.
- 米澤理可, 鮫島梓, 齋藤茂. 不正子宮出血に対する漢方薬の有効性についての検討. 産婦漢方研のあゆみ. 2018;35:32-6.
- 岩國亜紀子. 妊婦が自らに合ったつわり軽減方法を見出すセルフケア行動の向上を目指した看護援助プログラムによるセルフケア能力及びつわりへの効果検証. 日看科会誌. 2017;37:353-63.
- 穴戸恵理, 堀内成子. 無痛分娩の希望とその分娩転帰: 自然分娩との比較. 母性衛生. 2018;59:112-20.
- 飯田真理子, 堀内成子, 永森久美子. 妊婦健診の“女性を中心としたケア”の認識とマタニティブルーズおよび産後うつ病との関連. 母性衛生. 2018;59:81-9.
- 林育代, 鈴木麻希, 能瀬陽子, 湊聡美, 住友文, 二連木普輔 ほか. 日本人妊婦における妊娠前の体格, 体型認識と妊娠中の体重増加との関連. 肥満研究. 2017;23:233-40.
- 森本眞寿代, 南里美貴, 水田るり子, 川崎寿磨子. A大学病院で出産した高齢経産婦の産後1ヶ月までの育児不安の実態: 35歳以上と34歳以下の比較. 母性衛生. 2018;58:718-8.
- 青島友香, 山本記穂, 入駒慎吾, 村越毅. 無痛分娩群と非無痛分娩群の産褥初期の疲労感比較. 分娩と麻酔. 2017;99:70-4.
- 江川美保, 岡本和也, 西村史朋, 森野佐芳梨, 桑直人, 青山朋樹 ほか. 月経前症候群の管理におけるスマートフォンアプリを用いた症状記録システムの開発と臨床使用. 女性心身医学. 2016;21:105-13.
- 古井辰郎, 高井泰, 木村文則, 北島道夫, 中塚幹也, 森重健一郎 ほか. 本邦におけるAYA世代がん患者に対する妊孕性に関する支援体制: がん専門医調査の結果より. 癌と化学療法. 2018;45:841-6.

- 28) 井上理恵, 松田祐典, 長嶋小百合, 半田裕子, 曾我真弓, 田中基 ほか. 帝王切開後慢性痛の発生頻度とリスク因子の検討: 前向きコホート研究. 麻酔. 2018;67:800-8.
- 29) 齊藤裕也, 武山陽一, 石垣展子, 松浦類, 田邊康二郎, 吉永浩介. 産科危機的出血に対し施行した動脈塞栓術と子宮全摘術の比較検討. 産科と婦人科. 2018;85:983-6.
- 30) 黒川晴夫, 小島亜華里, 泰山裕. シンキングツール～考えることを教えたい～. [http://www.ks-lab.net/haruo/thinking\\_tool/short.pdf](http://www.ks-lab.net/haruo/thinking_tool/short.pdf) [accessed 2018-07-27]