

原 著 論 文

## クリティカルケア看護学研究における サンプルサイズとリッカート尺度の質問方式の関係

### The relationship between sample size of a test and the Likert scale levels in producing statistically meaning results for critical care research for nursing

井 上 正 隆 (Masataka F. Inoue)\* 山 田 覚 (Satoru Yamada)\*

#### 要 約

**目 的：**クリティカルケア看護学研究において、中小規模な試行を基にケアの効果検証を行う際の統計学的方略についての示唆を得るために、サンプルサイズとリッカート尺度の質問方式を検討した。

**方 法：**研究は、コンピューターでシミュレーションしたデータを用いてすべての分析を行った。1つの正規乱数を基に2件、3件、4件、5件、6件、7件、11件、21件の各リッカート尺度に相当するスコア化を行った。このデータを基に正規性の確認と差の検定を分析データ数を変化させて行った。

**結果及び考察：**6件法以下の段階化を行ったリッカート尺度では正規性は求められなかった。また、リッカート尺度を用いる場合、サンプルサイズは75件以上必要である。1群のサンプル数が75件程度の場合、t検定とMann-Whitneyのu検定の検出力の違いはなかった。このため6段階以下のリッカート尺度も、V A S法の簡便法として活用できる。

#### Abstract

**Purpose :** This research explores the relationship between sample size of a test and the Likert scale levels in producing statistically meaning results for critical care research for nursing.

**Method :** The research used computer simulated data. We divided each item into 2 level, 3 level, 4 level, 5 level, 6 level, 7 level, 11 level, and 21 level to test the best level size for the given sample size.

**Result and discussion :** The Likert scale which used below a 6 level method did not result in normal distribution. When using the Likert scale method, 75 or more sample sizes are required. When the number of samples is about 75, there is no difference in power of statistical test between t-test and u-verification.

キーワード：サンプルサイズ クリティカルケア看護 リッカート尺度 数的シミュレーション

#### I. は じ め に

クリティカルケア看護学分野における量的研究方法を実施するに際しては、他の看護学分野と比較し、下記に述べるように対象の特性と母集団規模に起因した特性があり、研究実施時は特に注意が必要である。

まず、クリティカルケア看護学の特徴を考慮すると、山勢ら<sup>1)</sup>が「クリティカルケアとは、突発的な事故や重篤な疾病、身体的侵襲の大きい手術などによって重要生体機能に大きな障害がもたらされ、生命の危機的状況に陥っている

対象に、精力的かつ集中的なケアを施す医療を指している」と定義するように、ケアの対象は容易に日常的なコミュニケーションが図れない状況であることが多い。このため、研究協力者が患者やその家族の場合は、他の研究分野に増して研究協力による負担と正確かつ容易に回答できるような研究方法を考慮する必要がある。つまり、統計学的正確さを求めすぎ、難解な質問紙や調査方法を採用した場合、研究協力者の研究協力による過大な負担を強いることや有効回答数の低下を来すことが危惧される。

\*高知県立大学看護学部

次に、どのような研究分野でも量的データを用いた研究をする場合、理想的な研究の工程では、データを収集する前に必要サンプルサイズを算出することが求められる<sup>2)3)</sup>。しかしながらクリティカルケア看護学の分野で量的研究を行う場合、倫理的問題や対象数の少なさから算出されたサンプルサイズを満たすことは、容易ではない。例えば、国内の高度救命救急センターは30施設、ドクターヘリ基地施設41施設であり、研究テーマにもよるが、元の母集団が少ない研究分野であると言える。このため、本研究分野で量的研究を行う研究者は、より検出力の高い方法でのデータ収集方法を採用する必要が他分野に増して必要である。

また2000年から2004年の発表された国内のクリティカルケア看護学研究に関する調査では、75%が量的研究デザインを採用し<sup>4)</sup>、本分野の研究の質を向上させるためにも対象の特性と母集団規模に起因した特性を考慮した研究手法の普及が不可欠である。

これらの現状を踏まえ、特にケアの効果検証を行うような場面を想定し、クリティカルケア看護学研究におけるサンプルサイズと質問紙法を用いる際の質問方式の検討を統計学的視点から行うことを目的に研究を実施した。これを基に以下のリサーチクエスチョンを策定した。

1. データが、正規分布を示す条件に関与する段階化(※1)とサンプルサイズの関係性を明らかにする。
2. サンプルサイズと段階化による検出力の関係を明らかにする。

(※1) 段階化：リッカート尺度を何段階の質問方式に設定するかのこと。段階化という用語は、一般的な用語ではないが、「質問方式」では「段階化」以外の内容を含み、他に適当な用語が無いので本文中ではこの用語を用いる。

また、これらの疑問を解明する中で、中小規模な試行を基にケアの効果検証を行う際の統計学的方略についての示唆を得る。

## II. 方 法

本研究で用いたデータは、コンピューターでシミュレーションしたデータを用いてすべての分析

を行った。データ生成の手順と分析の方法は、以下のとおりである。なお、シミュレーションには、Microsoft Excel、統計分析には、SPSSを用いた。

### 1. データの生成

#### 実験1：リッカート尺度の正規性の確認

##### 1) 正規性を持つ乱数の生成

平均0、標準偏差1の正規分布に対応する乱数を【=NORMINV(RAND(), 0, 1)】のコマンドを用いて発生させた。このデータは計量値であり、研究協力者が質問項目に対して持つイメージに相当する。

##### 2) 正規乱数のスコア化(※2)

上記の正規乱数をリッカートの尺度に合わせて整数のスコア化を行った。スコアは、図1のように $\pm 2\sigma$ の範囲に95%入るようにした。段階化は、2件、3件、4件、5件、6件、7件、11件、21件とした。スコア化したデータは、1つの正規乱数を基にそれぞれ行った。

(※2) スコア化：連続値の基データをリッカート尺度に合わせて整数化すること。リッカート尺度を用いた質問紙では、研究協力者が持つイメージは、スコア化して回答される。

##### 3) データベース作成

上記の方法を用い偶然の要素を排除するために50,000件のデータを生成し、その中から無作為にデータを抽出し、データ数がそれぞれn=1000、n=500件、n=250件、n=200件、n=150件、n=100件、n=75件、n=50件、n=40件、n=30件のデータベースを作成した。また、分析結果の偶発性を避けるため、データベースをそれぞれ10セット作成した。

### 2. 統計分析

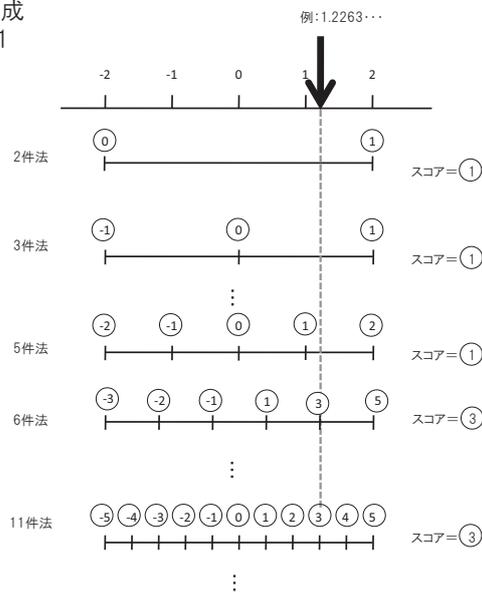
#### 1) 正規性の検証

先のプロセスで生成したデータベースごとに2件、3件、4件、5件、6件、7件、11件、21件にスコア化したデータに対し、Shapiro-Wilk検定を行い、正規性を確認した。既述のようにデータベースは、各10セット有るので、例えば4件法に関しては、Shapiro-Wilk検定を10回異なるデータで行っている。

# 実験1

1) 正規性を持つ乱数の生成  
平均=0、標準偏差=1

2) スコア化  
※ スコア化したものは、①のように数字に○をつけて表す



3) 50,000セットのデータを生成

4) 無作為にデータセットを抽出

5) データ数がそれぞれn=1000、n=500件、n=250件、n=200件、n=150件、n=100件、n=75件、n=50件、n=40件、n=30件のデータベースを作成した。

4)、5)の作業を10回繰り返した。

正規乱数及び2件、3件、4件、5件、6件、7件、11件、21件にスコア化したデータを1セットとする

# 実験2

1) 正規性を持つ乱数の生成 各郡ごとに50,000件の正規乱数を生成

データA群

平均=0、標準偏差=1

データB群

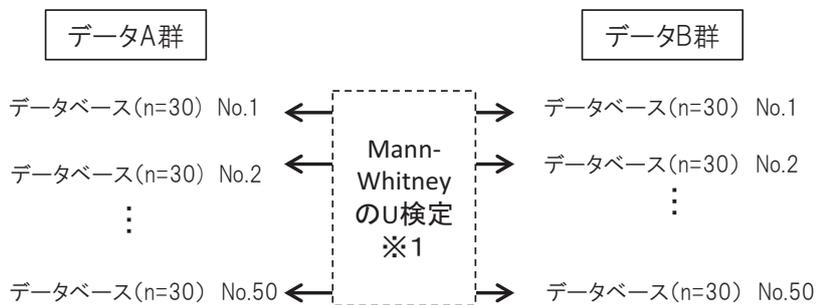
平均=0.5、標準偏差=1

2) スコア化 (実験1と同手順)

3) 無作為抽出

乱数を用い、データA群、データB群ごとにデータを抽出し、n=75件、n=50件、n=30件のデータベースを各50セット作成した

4) 平均の差の比較



5) 検出力の算出

※1: Mann-WhitneyのU検定は、No.1~No.50のデータベースごとに50回行い、p値の平均値を算出した。また算出されたp値が $p < 0.05$ となっているかを確認した。

検出力 = 「有意差あり」となった試行数 ÷ 50 (全試行数)

図1 研究方法

## 2) 分析

Shapiro-Wilk検定の帰無仮説は、標本が正規母集団からサンプリングされたものであるので、有意確率が0.05以上のものは正規分布をなしていると判別した。次に算出された10検定の有意確率の平均と正規分布と判別された割合を算出した。

### 実験2：リッカート尺度の段階化による検出力の差の検証

#### 1) 正規性を持つ乱数の生成

平均0、標準偏差1の正規分布に対応する乱数を【=NORMINV(RAND(), 0, 1)】のコマンドを用いて発生させた（以下データA群）。また同じく、平均0.5、標準偏差1の正規分布に対応する乱数を【=NORMINV(RAND(), 0.5, 1)】のコマンドを用いて発生させた（以下データB群）。なお、正規乱数での平均値0.5の違いは、スコア化した際に5件法での0.5スコア分に相当する。

#### 2) 正規乱数のスコア化

実験1と同様の手順でスコア化を行った。

#### 3) データベース作成

上記の方法を用い、データA群、データB群ごとに偶然の要素を排除するために50,000件の正規乱数を生成し、その正規乱数を基にスコア化を行った。さらにその中から無作為にデータを抽出し、データ数がn=75件、n=50件、n=30件（※3）のデータベースをデータA群、データB群ごとに作成した。また、分析結果の偶然性を避けるため、n=75件、n=50件、n=30件のデータベースをそれぞれ50セット作成した。

（※3）n=25件のデータベースは、データ数が少なすぎると考えたので、最小サイズのデータベースは、n=30件とした。

## 2. 統計分析

### 1) データA群とデータB群の差の確認と的中率の算出

正規乱数及び2件、3件、4件、5件、6件、7件、11件、21件にスコア化したデータについて、データA群とデータB群の差を確認するためにMann-WhitneyのU検定をn=75件、n=50件、n=30件の各50セットのデータベースごとに行った。次に、各検定で算出されたp値の平均と標準偏差、標準誤差を算出した。また本実験では、恣意的に平均

値の違うデータA群とデータB群間で検定を行っているので、理論上は検定で有意な差が認められるはずである。言い換えると検定で「有意差なし」とされた検定は、 $\beta$ エラーを示していると言える。これを検証するために、 $P < 0.05$ を有意水準とし、検定結果に有意差があるかを判別し、検出力(power)を求めた。

### 2) データサイズと段階化による検出力の関係の分析

データサイズと段階化が、検出力に対してどの程度の影響を示すのかを明らかにするために、前の分析で算出された有意確率を従属変数にし、データ数と段階化を独立変数にして重回帰分析をステップワイズ法を用いて行った。

### 3) データサイズによるパラメトリック検定とノンパラメトリック検定の検出力差の分析

生成した正規乱数をデータA群とデータB群間でt検定とMann-Whitneyのu検定をそれぞれ行い、的中率を算出した。この工程をn=30、n=50、n=75のデータベースごとに行った。さらに、t検定とMann-Whitneyのu検定のそれぞれの的中率をデータベースごとにクロス集計分析を行った。

### 4) t検定とMann-Whitneyのu検定の検出力の違いの判別

上記3)で算出した正規乱数に対するデータA群とデータB群間のt検定的中率と上記1)で算出したリッカート尺度に沿ってスコア化したデータに対するデータA群とデータB群間のMann-Whitneyのu検定的中率をデータベースごとにクロス集計分析を行った。この分析は、4件、5件、6件、7件、11件、21件に段階化したデータごとに行い、検定の6回の繰り返しとなるので、対馬<sup>5)</sup>が説明するBonferroniの不等式に基づく多重比較法に従い、この分析のみ有意水準を $P < 0.00625$  ( $0.05 \div 6$ )とした。

## III. 結果

### 実験：1

#### 1. Shapiro-Wilk検定の検定結果を表1に示す。

すべてのデータベースで正規乱数は、正規分布に基づいていると判別された ( $P < 0.05$ )。段階化に注目すると、正規性が確認できたのは、

7件法ではn=40以下のデータベースで、11件法ではn=100以下のデータベースで、21件法ではn=150以下のデータベースであった。また、50%以上の確率で正規性が確認できたのは、11件法ではn=50以下のデータベース、21件法ではn=100以下のデータベースであった。7件法の

データベースでは、50%以上の確率で正規性を確認することはできなかった。

一方、その他の2件法、3件法、4件法、5件法、6件法については、どのデータベースでも正規性を確認できなかった。

表1 生成データの正規性の確認

n=1000									
生成乱数	2件法	3件法	4件法	5件法	6件法	7件法	11件法	21件法	
0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.89	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
確率の平均	0.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

n=500									
生成乱数	2件法	3件法	4件法	5件法	6件法	7件法	11件法	21件法	
0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.64	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.86	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
確率の平均	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

n=250									
生成乱数	2件法	3件法	4件法	5件法	6件法	7件法	11件法	21件法	
0.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	
0.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
確率の平均	0.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

n=200									
生成乱数	2件法	3件法	4件法	5件法	6件法	7件法	11件法	21件法	
0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	
0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	
0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	
0.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	
0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
確率の平均	0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01

n=150									
生成乱数	2件法	3件法	4件法	5件法	6件法	7件法	11件法	21件法	
0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02
0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
0.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05
0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
確率の平均	0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02

n=100									
生成乱数	2件法	3件法	4件法	5件法	6件法	7件法	11件法	21件法	
0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.13	
0.42	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	
0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.08	
0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.07	
0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04	
0.54	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.17	
0.66	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.12	
0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	
0.76	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.06	
0.62	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	
確率の平均	0.57	0.00	0.00	0.00	0.00	0.00	0.02	0.07	

n=75									
生成乱数	2件法	3件法	4件法	5件法	6件法	7件法	11件法	21件法	
0.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.27
0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.15
0.95	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.13	0.11
0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.07
0.63	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.34
0.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.45
0.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.10
0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.10
0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.03
確率の平均	0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.16

n=50									
生成乱数	2件法	3件法	4件法	5件法	6件法	7件法	11件法	21件法	
0.33	0.00	0.00	0.00	0.00	0.00	0.01	0.11	0.11	
0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.02
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.15
0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.15	0.04
0.36	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.09	0.18
0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.32
0.81	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.28	0.44
0.81	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.26	0.36
0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.14
0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.02
確率の平均	0.35	0.00	0.00	0.00	0.00	0.00	0.01	0.11	0.18

n=40									
生成乱数	2件法	3件法	4件法	5件法	6件法	7件法	11件法	21件法	
0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.06	0.07
0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.10	0.04
0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.51
0.30	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.19	0.05
0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.07	0.08
0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.03	0.32
0.52	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.29	0.36
0.83	0.00	0.00	0.00	0.00	0.00	0.02	0.09	0.29	0.56
0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.07
0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.02
確率の平均	0.25	0.00	0.00	0.00	0.00	0.00	0.03	0.13	0.21

n=30									
生成乱数	2件法	3件法	4件法	5件法	6件法	7件法	11件法	21件法	
0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.07	0.07
0.17	0.00	0.00	0.00	0.00	0.01	0.02	0.17	0.08	0.12
0.06	0.00	0.00	0.00	0.00	0.00	0.03	0.02	0.41	0.73
0.62	0.00	0.00	0.00	0.00	0.00	0.01	0.06	0.62	0.12
0.16	0.00	0.00	0.00	0.01	0.02	0.06	0.17	0.17	
0.30	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.05	0.39
0.57	0.00	0.00	0.00	0.01	0.00	0.04	0.47	0.39	
0.20	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.08	0.16
0.17	0.00	0.00	0.00	0.01	0.02	0.01	0.10	0.12	
0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.16	0.06
確率の平均	0.26	0.00	0.00	0.00	0.00	0.01	0.04	0.22	0.23

表中の数字は、Shapiro-Wilk検定によって算出された有意確立を示す。

下段の「確率の平均」は、10回の試行の平均を示す。

## 実験：2

1) データ A 群とデータ B 群の差を確認的的中率の算出

分析データ数と段階化による Mann-Whitney の U 検定による p 値の変化を図 2 に示す。n=30、n=50 のデータベースでは、すべての段階化で p 値の平均は有意水準の 0.05 以上となった。n=75 のデータベースでは、4 件以上に段階化の多いデータで、p 値の平均値が 0.05 未満となった。

次に、分析データ数と段階化による Mann-Whitney の U 検定による検出率の変化を図 3 に示す。検出率は、n=75 のデータベースでは、すべての段階で 60% 以上となり、4 件法以上の段階化を行ったデータでは、70% 以上となった。一方、n=30 のデータベースでは、5 件法以上でも的中率は、50% 前後で推移した。

2) データサイズと段階化による検出力の関係の分析

有意確率を従属変数にし、データ数と段階化

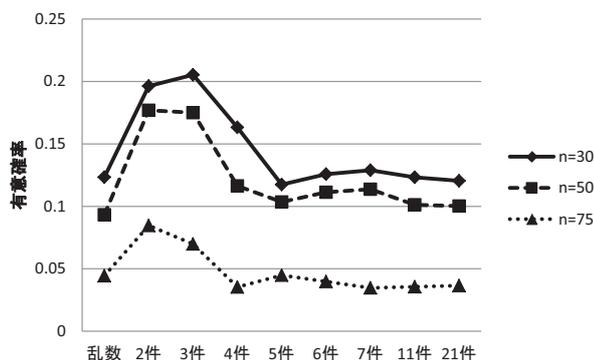


図 2 データ数と段階化による Mann-Whitney の U 検定の p 値の変化

を独立変数にして重回帰分析をステップワイズ法を用いて行った結果を表 2 に示す。なお、従属変数となる有意確率は、正規分布と判別された ( $P=0.089$ )。分析の結果、 $R^2=0.754$ 、調整済み  $R^2=0.731$ 、分散分析 P 値  $<0.001$  であり、データ数 (標準化係数  $\beta=-0.813$   $P<0.001$ ) と段階化 (標準化係数  $\beta=-0.306$   $P<0.001$ ) の両変数がモデルに採択された。共線性は、認めなかった ( $VIF=1.0$ )。

3) データサイズによるパラメトリック検定とノンパラメトリック検定の検出力差の分析

データ A 群とデータ B 群のそれぞれの生成乱数の比較を t 検定と Mann-Whitney の u 検定で行い、的中率をデータ数が n=30、n=50、n=75 のそれぞれのデータベースごとにクロス集計分析を行った結果、どのデータベースでも t 検定と Mann-Whitney の u 検定の的中率に有意な差はなかった (n=30 :  $P=0.841$ 、n=50 :  $P=0.829$ 、n=75 :  $P=0.074$ )。

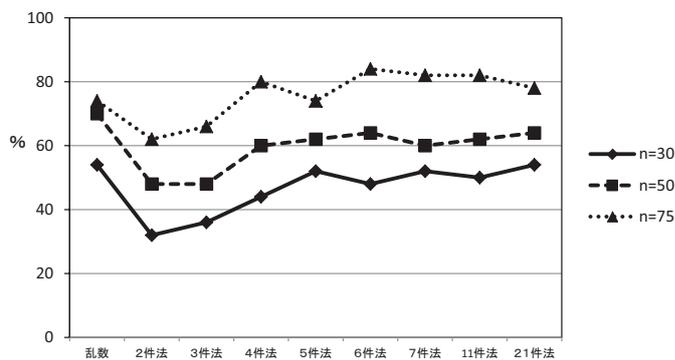


図 3 データ数と段階化による Mann-Whitney の U 検定の検出率変化

表 2 有意確率を従属変数にし、データ数と段階化を独立変数にした重回帰分

	偏回帰係数	有意確率	標準化係数	VIF
データ数	-.002	.000	-.813	1.000
段階化	-.003	.010	-.306	1.000
(定数)	.243	.000		

$R^2=0.754$  調整済み  $R^2=0.731$   
分散分析 P 値  $<0.001$

#### IV. 考 察

##### 1. データが、正規分布を示す条件への段階化とサンプルサイズの関係

例えば、平均の差を検定する場合では、独立したサンプルの t 検定と Mann-Whitney の U 検定のようにパラメトリック検定に対応したノンパラメトリック検定が考案されている場合が多い。厳密には、独立したサンプルの t 検定は平均を比較し、Mann-Whitney の U 検定は分布を比較しているが、この違いを考慮しなければ、データの分布が正規分布に従うかを考慮する必要はあまりない。しかし、例えば重回帰分析を行う場合は、従属変数が正規分布に従う必要があり、データ分布が正規分布に従うかは、重要な問題となる。

例えば質問項目 A から D までの合計点を算出するような複数の質問項目を加法して従属変数として分析に用いる場合について考えてみると、50%以上の確率で正規性が確認できるのは、段階化の合計が11段階以上でサンプル数が50件程度の条件である。クリティカルケア看護学研究でよく用いられている4件法や5件法であれば、3題の質問項目の得点を加法し、サンプルサイズが $n=50$ 以上あればデータの正規性が期待できる。別の言い方をすると、研究協力者にとって難解な長い文章の質問項目を立案するよりは、平易な文章の質問項目を複数立案した方が、統計的には有利であると言える。また、変数間の影響度の差異に注意し、少なくとも抽象度の統一を行う必要があると言える。

では、単独の質問項目を従属変数として用いたい場合は、どのような工夫をすればよいだろうか。このようなケースとしては、疼痛や倦怠感などの研究協力者が感じる症状を研究する場合を想定した。先に述べたように正規性の問題から、1つの質問項目で従属変数になるには、段階化が11段階以上の尺度を用いた質問項目が必要である。

クリティカルケア看護領域の患者特性について考えてみると、Newcomerらは、ICU入室中患者においてストレス物質の糖質コルチコイドと記憶力に関係があるとし<sup>6)</sup>、McEwenらもストレスと認知障害に関係があるとしてしている<sup>7)</sup>。

また、福田らは<sup>8)</sup> ICU入室経験のある患者の62.5%で記憶の欠落か非現実的な体験をしているとしており、クリティカルケア看護領域で患者を対象とした研究では、11段階の段階を研究協力者が判別できるのかという問題が予想される。このため、実際に研究を行う際の工夫としては、VAS法を用いて「昨日と比べて倦怠感ほどの程度か？」などを質問し、複数の日程で同一の質問紙を用い、過去に研究協力者が回答したスコアからの距離を数値化するのが適当ではないかと考えた。

研究目的から逸れる内容になるが、本研究結果からはShapiro-Wilk検定をリッカート尺度の正規性の確認に用いる場合に条件制限を考慮すべきとの示唆も得るものであった。中心局限定理では、サンプル数が多くなると正規分布に近づくと言われており、計量値の性質を持つ正規乱数は、サンプルサイズが大きくなるに従ってShapiro-Wilk検定で算出されるp値は、上昇する傾向にあり、中心局限定理を裏付けるものであった。一方、リッカート尺度を想定しスコア化したデータでは、最も段階化の細かい21件法でも $n=150$ 以上の条件では、正規性を認めなかった。Shapiro-Wilk検定の有用性について議論するのは、本研究目的から外れたものであり、本検定法を正規性の確認に用いている本研究では、自己矛盾に陥る内容であるが、今後検証が必要であると考えた。特に、リッカート尺度を用いて、概念の構造分析を行う際は、本研究とは別の議論が必要になってくると考える。

##### 2. データサイズと段階化による検出力の関係

本研究結果からは、対応の無いサンプルの検定を行う際は、サンプルサイズを各群75件以上用いると、4件以上の段階化を行ったリッカート尺度では、70%以上の確率で標準偏差0.5の距離に相当する平均の差を検出することができる。また、サンプルサイズが50件の場合も同様に4件以上の尺度で60%前後の的中率を得ることができる。また、サンプルサイズが30件になった場合は、5件以上の尺度で50%前後の検出力を得ることができ、4件法の使用は適切ではない。医療統計では、検出力を0.8から0.9に設定することが適当とされており<sup>9)</sup>この

基準 (power>0.8) に適合するのは、サンプルサイズが75件以上の6件法、7件法、11件法である。

また、データサイズと段階化による検出力の関係について明らかにするために行った重回帰分析結果からは、有意確率に与えるデータ数の影響力は段階化の約2.7倍あり、データ数が同じ場合は段階化が多い方が有意確率を下げると言える。しかし、4件法と5件法の比較では、有意確率が0.05未満となるのは、計算上それぞれn=91件とn=90件であり、大きな差はない。なお、11件法と21件法では有意確率が0.05未満となるのは、それぞれn=81件、n=66件である。

クリティカルケア看護学会誌に2009年から2013年に投稿された患者もしくは家族を対象とした研究は5件あるが、その中で文部科学省と厚生労働省が示す疫学研究に関する倫理指針において観察研究に該当すると考えられる2件<sup>10)11)</sup>を除くと、研究協力者数は44人<sup>12)</sup>、21人<sup>13)</sup>、71人<sup>14)</sup>とサンプルサイズは小規模であった。サンプルサイズが小さいことが予測されるクリティカルケア看護学分野の研究では、本研究結果を踏まえると、リッカート尺度の段階化を多くするか、V A S法を用いることが適当であると言える。

これらの考察をまとめると、リッカート尺度を用いる場合、サンプルサイズは少なくとも比較する1群のサンプル数が75件以上は必要であり、6件以上に段階化したリッカート尺度かV A S法を用いることが必要であると言える。クリティカルケア看護学分野の研究を想定すると、リッカート尺度の段階化を多くすることが研究対象者にとって難しい場合は、本調査の前にどの程度段階化が行えるかを調査し、リッカート尺度の段階化を決定することが必須になる。また、両群の差が標準偏差0.5の距離以上の隔たりを持つ指標を選択することが有用である。

ところで、t検定とMann-Whitneyのu検定の検出力の違いについては、データ数がn=30、n=50、n=75の各データベースでt検定とMann-Whitneyのu検定の的中率に有意な差は無く、データ数が各群75件未満の条件では、検出力に差が無いと結論付けた。ただ、データ数が増えるに従い、P値は低下する傾向にあるので、データ数が75件以上増えるとt検定とU検定の

検出力に違いが出ると推測される。クリティカルケア看護学分野の研究での応用を考慮すると、1群のサンプルサイズが75件以下と予測される場合は、検出力の差に注目して正規分布が期待される段階化を行ったリッカート尺度を立案する必要はないと言える。このため、本研究結果で正規性が認められなかった6段階以下に段階化したリッカート尺度でも、研究協力者の特性を考慮し、V A S法よりも研究上有利と判断した場合は、V A S法の簡便法として活用できると言える。

### 3. 本研究結果から得られたクリティカルケア看護学における量的研究への示唆

日本学術会議の高度実践看護師に関する提言<sup>15)</sup>では、高度実践看護師を「個人、家族、及び集団に対して、ケアとキュアの融合による高度な看護学の知識／技術を駆使して、対象の治療・療養過程の全般を管理・実践することができる看護師」と定義しており、今以上に看護師の実践基盤となる根拠を示す研究が重要になってくる。また、井上<sup>16)</sup>は、看護師の役割拡大とクリティカルケア領域の未来像について「医行為を単なる行為の単位として捉えるのではなく、それを看護ケアと融合させることでその効果が高まることが期待される」と述べ、看護ケアとキュアの融合によって看護の効果向上を目指していく必要があるとしている。これらの提言を実現する方策の一つとして、ケアの効果検証を行う量的研究が必要である。新たな介入方法の検証としては、R C T (Randomized Controlled Trial) が医学や薬学分野で多用されている。高橋<sup>17)</sup>は、R C Tを看護学で行う場合の課題として、看護介入方法の再現性の確保と並び、研究協力者の確保とアウトカム指標として用いる心理社会的な測定指標の信頼性と妥当性の確保を挙げている。本研究結果からこの課題への示唆としては、V A S法が第一選択として考えられるが、プレテストにおいて患者特性を考慮し、6件法を中心にして段階化を調整したリッカート尺度を用いることが有用であることが挙げられる。またデータ数に関しては、1群に75件程度確保できるとデータが安定化することが挙げられる。

## V. 結 論

リッカート尺度の正規性については、6件法以下の段階化を行った尺度では正規性は求められず、1000件程度のサンプルサイズがある場合でも、正規分布であることが前提となる検定の使用は難しい。

クリティカルケア看護学分野でケアの効果測定を計画し、リッカート尺度を用いる場合、標準偏差0.5の距離に相当する平均の差を検出するためには、サンプルサイズは少なくとも比較する1群のサンプル数が75件以上は必要であり、6件以上に段階化したリッカート尺度かVAS法を用いることが必要である。また、両群の差が標準偏差0.5以上の距離を確保できるようなアウトカム指標を選択すべきである。

1群のサンプル数が75件程度の場合、t検定とMann-Whitneyのu検定の検出力の違いはなく、正規性が期待できない6段階以下に段階化したリッカート尺度でも、クリティカルケア看護学分野の研究で予想される研究協力者の認知力の低下等を考慮し、VAS法よりも研究上有利と判断した場合は、VAS法の簡便法として活用できる。

## &lt;引用・参考文献&gt;

- 1) 山勢善江、山勢博彰：クリティカルケア看護の概念（成人看護学B急性期にある患者の看護Iより（氏家幸子監修、泉キヨ子、大森武子編）、廣川書店、41-47、2001.
- 2) 森實敏夫：わかりやすい医療統計学、メディカルトリビューン、31、2004.
- 3) 永田靖：サンプルサイズの決め方、朝倉書店、1-15、2003.
- 4) 山勢博彰：わが国のクリティカルケア看護に関する研究の動向、看護研究、38(2)、89-99、2005.
- 5) 対馬栄輝：SPSSで学ぶ医療系データ解析、東京図書、152-153、2007.
- 6) Newcomer JW, Craft S, Hershey T, Askins K et al: Glucocorticoid-induced impairment in declarative memory performance in adult humans, J Neurosci, 14, 2047-2053, 1994.
- 7) McEwen BS, Sapolsky RM: Stress and cognitive

function, Curr Opin Neurobiol, 5, 205-216, 1995.

- 8) 福田和明、黒田裕子：重症患者家族ニーズに対するクリティカルケア看護師の認知構造モデルの構築、日本クリティカルケア看護学会誌、8(1)、17-28、2012.
- 9) 森實敏夫：わかりやすい医療統計学、メディカルトリビューン、32、2004.
- 10) 梶原弘平、小島昌人、古賀雄二、清田亜紀：緊急入院した脳神経疾患患者の特性と入院期間との関連、日本クリティカルケア看護学会誌、7(3)、35-42、2011.
- 11) 濱崎真由美、平松八重子、秋山直美：A大学病院ICU摂食・嚥下機能評価表の開発と有用性の検討、日本クリティカルケア看護学会誌、7(3)、26-34、2011.
- 12) 若崎淳子、谷口敏代、掛橋千賀子、森將晏：成人期初発乳がん患者のQOLに関する縦断研究（その1）—手術前から手術後1年までのQOLの経時変化とその要因—、日本クリティカルケア看護学会誌、6(1)、1-15、2010.
- 13) 益田美津美：血管内手術を選択した未破裂脳動脈瘤患者が抱く不確かさの構造と経時的变化に基づく看護支援の検討、日本クリティカルケア看護学会誌、8(3)、1-14、2012.
- 14) 木下佳子：記憶のゆがみをもつICU退室後患者への看護支援プログラム開発とその有効性に関する研究、日本クリティカルケア看護学会誌、7(1)、20-35、2011.
- 15) 南裕子、太田喜久子、内布敦子ら：提言「高度実践看護師制度の確立に向けて—グローバルスタンダードからの提言—」、日本学術会議健康・生活科学委員会看護学分科会、<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-21-t135-2.pdf>、2011
- 16) 井上智子：看護師の役割拡大とクリティカルケア領域での未来像—特定看護師（仮称）創設の動きの中で—、日本クリティカルケア看護学会誌、7(1)、1-7、2011.
- 17) 高橋 さつき：わが国の臨床看護技術に関するランダム化比較試験のシステムティックレビュー、群馬県立県民健康科学大学紀要、5、57-72、2010.