

20031061

厚生労働科学研究研究費補助金

医療技術評価総合研究事業

# 歯科技工士資格試験における技術評価等に関する研究

平成15年度 総括研究報告書

主任研究者 末瀬 一彦

大阪歯科大学歯科技工士専門学校

平成16(2004)年3月

# 目次

## I. 総括研究報告

歯科技工士資格試験における技術評価等に関する研究

末瀬一彦

1 - 48

研究報告書

歯科技工士資格試験における技術評価等に関する研究

主任研究者 末瀬一彦 大阪歯科大学歯科技工士専門学校校長

**研究要旨** 近年、高度化・多様化する歯科医療界にあって、社会や時代の要請に適応した質の高い歯科技工士を確保し、品質管理された歯科技工物を安定的に供給できることを目的に歯科技工技術教育を行なわなければならない。歯科技工士養成施設において指定規則に基づく2年ないしは3年の教育カリキュラムを修了したのち、歯科技工士法第11条によって歯科技工士として必要な知識および技能について判定する資格試験を受験し、合格しなければならない。現在、歯科技工士資格試験は学説試験および実技試験が課せられているが、実技試験は重要な技能評価法であり、他の医療関係職種における試験制度の改善などを鑑みた場合、情報公開の流れなども受け、国家資格の試験としてふさわしい客観的評価基準が担保された試験でなければならない。そこで、本研究においては歯科技工士資格試験の実技試験の指標開発を行うとともに、具体的な実施方法や採点基準、評価方法などについて詳細な検討を行い、実技試験の実施手法の確立を目指す。本研究を実施するにあたっては、従来から各都道府県に委託、実施されている出題課題を十分検討かつ反映し、義歯人工歯排列、歯冠蝕形成、鑄造鉤蝕形成、線鉤屈曲の4課題を1つの模型にまとめた試験用模型を策定し、全国歯科技工士教育協議会の協力のもと4校、120名の学生を対象に実技試験を実施した。実技試験の評価は「妥当性」、「信頼性」、「公平性」、「効率性」および「経済性」を考慮して所属の異なる評価者を選定し、大学教員、歯科技工士専門学校教員、臨床歯科医師および臨床歯科技工士の4群からなる28名の評価者によって、概略的評価および細分化評価について4段階的評価法による採点評価を行った。その評価結果をもとに、評価方法の関連性、評価者群間および評価者内の評価のばらつき、合格基準および実技試験評価結果と学内成績との関連性などについて検討した。さらに、今回実施した模擬的資格試験に関して、評価時間や試験時間、試験問題の内容や試験用模型について評価者および受験者からのアンケート調査も合わせて行った。

その結果、各評価者の総合評価合計値は正規性の仮定しうるデータであることが判明し、本研究の統計解析は正規性を前提とした手法を用いた。概略的評価と細分化評価との関係において、概して両者の評価結果は一致していた。特定の評価者群においてはすべての評価項目においてバラツキが少なく、均等に評価されたが、評価者群によっては概略的評価がある特定の評価項目に依存する傾向があった。試験課題によっては評価者群内でのバラツキが大きく、評価者群間においても有意の差が認められた。繰り返し評価による評価者個人内の再現性は低かった。総合評価と学内における成績評価を比較したところ、成績上位者の判別は可能であるが、成績中位者および下位者については評価者間での評価程度が異なり、判別不可能であった。概略的評価および細分化評価に要した評価時間は評価者群間において異なった。また試験直後に行った評価者および受験者に対するアンケート調査では試験実施時間について60%の受験者が適切であるとし、評価者の概略的評価では時間がかからなかったが、細分化評価ではかなり時間を要した。試験問題の内容については評価者の75%は適切であるとしたが、試験用模型については評価者および受験者から改善点を指摘した。

## 分担研究者

西田紘一 日本歯科大学歯学部

佐藤温重 明倫短期大学

田上順次 東京医科歯科大学大学院

鳥山佳則 茨城県保健福祉部

## A. 研究目的

歯科技工士資格試験は、厚生労働大臣の委託によって歯科技工士養成施設の所在地の都道府県知事が毎年 1 回実施しているのが現状であり、平成 13 年度厚生労働科学研究「今後の歯科技工士の養成方策等に関する総合的研究」で行った調査によれば試験実施時期、受験料、出題数、出題形式などが異なり、さらには実地試験における判定基準の統一の見解がないのが実態である。日本全国いかなる場所、いかなる時においても国民に対して良質な歯科医療サービスを安定的かつ効率的に提供するためには歯科技工士資格試験の全国統一化は必須で、その実施方式が適切なものでなければならない。特に歯科技工士資格試験における実技試験は重要な技能評価法であり、他の医療関係職種における試験制度の改善などを鑑みた場合、情報公開に対する妥当性も考慮して、国家資格の試験としてふさわしい客観的評価基準が担保されたものでなければならない。さらに歯科技工士における実技試験を行うにあたっての具備条件として、歯科技工士として具備すべき固有の知識および技能としての「妥当性」、評価者間で評価基準の適用が異ならず評価の再現性、一致度が高い「信頼性」、試験内容に不公平さが生じない「公平性」、多くの受験生の作品が迅速に評価できる「効率性」、比較的大幅

な設備費用などを導入しなくても評価の効果が大きい「経済性」などが挙げられる。

そこで本研究においては、歯科技工士資格試験における実技試験の技術評価基準の確立をめざして、現在実施されている資格試験出題内容を基盤にして策定した試験用モデルを用いて模擬的資格試験を実施し、その結果について所属施設の異なる評価者群によって採点評価を行い、異なる評価法の相関性や評価者群間ならびに評価者内の比較検討、試験成績と学内成績との関係、合格基準の設定と不合格率などについて統計学的に分析を行った。さらに、模擬的資格試験の受験者および評価者に対して、試験時間あるいは評価時間、試験内容、試験用モデルについてアンケート調査を行った。

## B. 研究方法

現在実施されている歯科技工士資格試験の出題内容は総義歯排列（歯肉形成含めて）、歯型彫刻および任意問題（支台築造、ブリッジの支台装置、架工歯の蝸型採得、鑄造鉤の蝸型採得、線鉤の屈曲、部分床義歯大連結子の蝸型採得のなかから出題）によって行われている。本研究に用いた模擬的試験用のモデルは、これらの出題内容を基盤にして検討し、全国統一試験が実施された場合の実施時間を十分考慮し、さらに、試験のための準備が教員や受験生にとって負担過重にならないことを配慮した。また、平成 14 年度厚生労働科学研究「今後の歯科技工士養成方策等に関する総合的研究」において資格試験用モデルの妥当性について歯科医師会、歯科技工士会ならびに歯科技工士養成施設に対して行ったアンケート調査に

においても概ね賛同を得た。すなわち本研究で用いた試験用模型の内容および実施時間は表 1 に示すように部分床義歯の人工歯排列および歯肉形成 (90 分)、全部鑄造冠蟻形成、鑄造鉤蟻形成ならびに線鉤屈曲 (120 分) の 4 課題で、作業模型を簡易型咬合器に装着した状態 (図 1) で受験生に配布し、出題は図 2 に示す歯科技工指示書に基づいて試験を実施した。本研究を実施するにあたって模擬的資格試験の協力校は、全国歯科技工士教育協議会加盟の東京医科歯科大学歯学部附属歯科技工士学校 (20 名)、日本歯科大学附属歯科専門学校歯科技工士学科 (40 名)、明倫短期大学歯科技工士学科 (20 名) および大阪歯科大学歯科技工士専門学校 (40 名) の 4 校で、歯科技工士学科 (本科) 2 年生および歯科技工士専攻科 (実習科) 1 年の学生合計 120 名である。任意に抽出された学生について平素の学内成績 (A, B, C ランク表示) を付記し、性別、年齢を含めて受験者リストを作成した。模擬的資格試験はほぼ同一時期にそれぞれの学校において実施し、試験終了後直ちに大阪 (大阪歯科大学歯科技工士専門学校)、ついで東京 (日本歯科大学附属歯科専門学校) にすべての模型を集配し、評価者による採点評価を行った。

採点評価者は所属の異なる 4 群、すなわち大学教員 (補綴系教授・助教授・講師)、歯科技工士専門学校教員 (校長・教務主任)、臨床歯科医師 (開業歯科医師) および臨床歯科技工士 (開設歯科技工士) の 28 名が行った (表 2)。評価は概略的評価および細分化評価の順で 2 回行い、評価に要した時間を記載した。概略的評価はそれぞれの 4 課

題について 0 (25 点), 1 (50 点), 2 (75 点), 3 (100 点) の 4 段階的評価とし、0: 未完成、1: 完成度が低い、2: 普通、3: 完成度が高いによって採点評価した (図 3)。さらに、6 名の評価者においては日を異にする 2 回の概略的評価を行った。細分化評価にあたってはそれぞれの 4 課題について 5 つの評価項目を設定し (図 4)、それぞれの評価項目について採点前に着眼点を指示し (表 3)、1 課題 5 評価項目について先と同様の 4 段階的評価を行った。

本研究で行った模擬的資格試験の採点評価の妥当性を検討するために以下の項目について分析を試みた。

- ① 概略的評価総合計値の正規性
- ② 概略的評価と細分化評価の相関
- ③ 概略的評価と細分化評価の評価値における評価者群間の比較
- ④ 概略的評価と細分化評価の構造解析
- ⑤ 概略的評価における評価者個人内の分散度
- ⑥ 概略的評価と細分化評価の所要時間の比較
- ⑦ 概略的評価における総合計評価値と学内成績の比較
- ⑧ 概略的評価における評価値分布と合格率

さらに、今回実施した試験について、評価者および受験者に対して評価・採点時間または試験時間、試験問題の内容、試験用模型についてそれぞれアンケート調査を行った。

(倫理面への配慮)

本研究における倫理面での問題は少ないと考えるが、受験者および評価者に対して

は研究の主旨および目的の説明を十分行い、プライバシーの保護には十分配慮した。また、試行のための「資格試験の実施および評価」についてその目的を十分理解させ、決して強制的ではなく、参加しなくても不利益がないことを十分説明し、承諾が得られたものに対してのみ実施した。

## C. 研究結果

### 1) 評価結果について

#### ① 概略的評価総合計値の正規性

評価者 28 名による受験者 120 名の採点評価値は 0 (25 点)、1 (50 点)、2 (75 点)、3 (100 点) に点数変換したのち、それらの概略的評価総合計点数の分布を 4 群の評価者群ごとに検討した。概略的評価総合計値の正規確率プロット (図 5) はいずれの評価者群においても直線性を示した。すなわち評価者各群の総合計値はいずれも近似的に連続性正規分布標本として解析可能なデータ群であることが確認された。

#### ② 概略的評価と細分化評価の相関

本研究において採用した 2 つの評価法、すなわち概略的評価値と細分化評価値との関係を明らかにするために、概略的評価総合計値および細分化評価総合計値について試験課題別に、評価者群ごとに  $n = \text{受験者 } 120 \text{ 名} \times \text{評価者数}$  の散布図を作成し、2 変量間の相関係数 ( $r$  値) を評価総合計値および各試験課題別評価値について評価者群ごとに求めた。評価者第 1, 2, 3, 4 群における概略的評価総合計値と細分化評価総合計値の散布図を図 6-1, 2, 3, 4 に示したが、それ

ぞれの  $r$  値は 0.937, 0.956, 0.930, 0.903 で、概略的評価総合計値と細分化評価総合計値は高い相関を示した。

4 つの試験課題別評価総合計値について評価者群ごとに求めた 2 変量間の相関係数 ( $r$  値) を表 4 に示すが、試験課題別および評価者群別にみても概略的評価値と細分化評価値は相関していることが明らかとなった。

#### ③ 概略的評価と細分化評価の評価値における評価者群間の比較

4 つの評価者群間の概略的評価および細分化評価の評価値の比較を評価平均値の差の検定から検討した。行、列とも要因によって分類に水準化がなされていないため、かつ 1 受験者ごとに全評価者 28 名の評価を比較する必要があることから、従属 2 標本の  $t$  検定を用いた。受験者数が 120 名と大標本であったため、データの検出力は極めて高かった。概略的評価に関しては、評価者第 1, 2, 3 群および 4 群の試験課題 1 の評価値はそれぞれ  $70.1 \pm 12.8$ 、 $72.9 \pm 12.9$ 、 $66.2 \pm 13.9$  および  $71.0 \pm 11.1$  であり、また試験課題 2, 3, 4 の評価値は表 5 に示すとおりであり、評価者第 3 群の評価値が常に他の群と相違していた。試験課題 1-4 の評価総合計値でみると、それぞれ  $278.5 \pm 40.2$ 、 $277.6 \pm 46.4$ 、 $256.8 \pm 45.9$  および  $277.5 \pm 36.7$  であり、評価者第 1, 2 および 4 群の評価総合計値が類似した値に収斂し、差が検出されないのに対して評価者第 3 群では他の 3 つの評価者群すべてとの間に差が検出された。

細分化評価に関しては、試験課題 1 の

評価者第 1, 2, 3 および 4 群の評価値はそれぞれ  $78.2 \pm 9.1$ 、 $78.7 \pm 10.4$ 、 $78.2 \pm 9.1$  および  $76.8 \pm 10.6$  であり、試験課題 2, 3 および 4 の評価値は表 6 に示すとおりであり、試験課題 1-3 においては評価者第 1 群と第 3 群、第 2 群と第 4 群が同様な評価をしていたことが明らかとなった。試験課題 1-4 の細分化評価総合計値でみるとそれぞれ  $305.0 \pm 33.1$ 、 $297.3 \pm 35.4$ 、 $302.7 \pm 37.0$  および  $300.6 \pm 34.9$  であり、細分化評価総合計値は発散しており、評価者すべての群間で 5%未満の棄却率で有意の差が検出された。

概略的評価および細分化評価の評価値における評価者群間の比較により、概略的評価においては評価者第 3 群以外の各群は類似した評価値であることが明らかとなったが、細分化評価において評価者群間で相違があることが明らかとなった。

各評価者群の概略的評価における精度を比較するために、各評価者群における全受験者について試験課題別の評価値の変動係数 (CV 値) を求め、CV 値を評価者群間で比較し、評価者群内の評価の一致度を検討した。表 7 に示すように試験課題 1 に関しては評価者第 3 群が極めて大きなバラツキを示した。また、試験課題 2 および 3 に関しては評価者第 3 群および第 4 群のバラツキが大きく、第 1 群および第 2 群に対して有意の差が検出された。試験課題 4 に関してはすべての評価者群間での相違が小さくよく一致した評価を示した。

概略的評価における各評価者群の評価精度は試験課題により異なっていたが、評価者第 1 群および第 2 群の評価は他の評価者群より良好な精度であることが明らかとなった。

#### ④ 概略的評価と細分化評価の構造解析

概略的評価と細分化評価には良好な相関関係が認められたが、さらに概略的評価のバラツキの原因を細分化評価と対比することによって構造的に明らかにすることを試みた。受験者 120 名の各試験課題の概略的評価と細分化評価との関係を、概略的評価を従属変数として評価者群それぞれについて重回帰分析を行った結果を図 7-1, 2, 3, 4 および表 8 に示す。なお、重回帰分析に先立って主成分分析により構造解析を試みたが、第 1 成分しか検出できず直線回帰と同様の構造であった。細分化評価の評価項目 1-5 まで同質の因子を評価していることが判明した。したがって結果がより明確に解釈される重回帰分析による構造解析を行った。

得られた細分化評価における評価項目ごとの標準回帰係数は、各試験課題の概略的評価における各評価項目の重要度を示している。評価者群第 2 群(図 7 の No. 10-18)は、すべて試験課題にわたってバラツキが少なく、さらに評価項目 1-5 のすべてを均等に考慮した評価が行われたことを示している。これに対して評価者第 3 群(図 7 の No. 19-23)および第 4 群(図 7 の No. 24-28)では概略的評価がある特定の評価項目に依存する傾向が強く、しかも依存する評価項目が

評価者個人によって異なることを示した。評価者第1群(図7のNo.1-9)では評価者第2群(図7のNo.10-18)に類似した評価を行っていた。

概略的評価と細分化評価の構造解析から各評価者群が概略的評価において細分化評価項目1-5の各々に依存する度合いに差があることが明らかとなった。

#### ⑤ 概略的評価における評価者個人内のばらつき

評価者個人内のばらつきを明らかにするために6名の評価者が概略的評価について、同一試験課題に対する2回の繰り返し評価を行った。従属2標本のt検定の結果を表9に示す。試験課題1においては、5名の評価者で2回の評価値に不一致が検出された。その他の試験課題については、2名の評価者では繰り返し評価値が一致していたが、その他の評価者では一致あるいは不一致があった。全体として繰り返し測定においては評価値が一致しにくいことを示していた。しかし、2回の評価における評価値の平均値の差の検定は最大値で12.7、評価段階1ランク未満であった。

したがって、概略的評価における評価者個人内のばらつきは認められるが、評価に大きな影響を与えるものではないと考える。

#### ⑥ 概略的評価と細分化評価の所要時間の比較

概略的評価および細分化評価に要した所要時間を表10にまとめた。評価者第1-4群の平均評価時間は、概略的評

価においては $92.1 \pm 12.9$ 時間、細分化評価においては $207.1 \pm 37.2$ 時間で、概略的評価の所要時間は細分化評価の約50%で、短時間での評価が可能であった。評価者群間の所要時間の比較はデータ数が少なく困難であったが、評価者第1群および第2群の所要時間は評価者第3群および第4群に比較して短時間であり、またバラツキは少ない傾向であった。

#### ⑦ 概略的評価における総合計評価値と学内成績の比較

今回の試験に受験した学生の概略的評価総合計値と平素の学内成績の関係を評価者群別にグラフで表したのが図8である。歯科技工士学科(本科)2年生Aランク(成績上位32名)の学生の評価値、Bランク(33名)およびCランク(31名)の学生の評価値は、評価者第1群の評価値ではそれぞれ $288.9 \pm 42.7$ 、 $260.2 \pm 50.5$  および  $268 \pm 54.5$  点で、Aランクの学生の評価値はBランクおよびCランクの学生より高点であったが、BランクとCランクとの評価値間には有意の差が認められなかった。評価者第2-4群の評価においても評価値には差があるが総合計評価値と学内成績の比較結果は評価者第1群と同様であった。歯科技工士専攻科(実習科)1年生Aランク(10名)、Bランク(6名)およびCランク(8名)の学生の評価値は、評価者第1群の評価ではそれぞれ $313.2 \pm 46.5$ 、 $285.8 \pm 55.4$ 、 $300.4 \pm 62.2$  点で、Aランクの学生の評価値はBおよびCランクの学生の評価値より高点であったが、AランクおよびBランクとCランクとの間に

は有意の差が認められなかった。また、評価者第3群および第4群の評価においてはAランク、BランクおよびCランクの評価値の間には差がみられなかった。歯科技工士学科2年生と歯科技工士専攻科1年生のそれぞれの受験者全体の評価値は、いずれの評価者群とも歯科技工士専攻科学生で高い評価値であった。評価者第1群による評価では歯科技工士学科2年生と歯科技工士専攻科1年生の受験者全体の平均はそれぞれ272.4±50.9、302.9±55.5点であった。

#### ⑧ 概略的評価における評価値分布と合格率

各試験課題別評価値が総合計評価値に占める位置を知るために、概略的評価における受験者の総合計評価値を各試験課題の評価点25点、50点、75点および100点について区分したときの延べ度数を調査した。その結果を図9の3次元ヒストグラムに示す。試験課題1-4ともに各試験課題で高得点(75点または100点)を得たものは、総合計評価値においても高得点を得ていたが、試験課題3および4では低得点(25点および50点)を得たものの度数が大きかった。評価点25点(不合格)を有する受験者の総合計評価点は課題1および3においては $\leq 100-300$ に分布していた。また、課題2および4においては $\leq 100-360$ に分布していた。25点の評価点を有する受験者といえども総合計評価点では高い得点を得ている例があることが明らかとなった。

合格基準として合格点を400点満点

の60%すなわち240点以上、および参考に200点以上としたときの不合格者数および不合格率を評価者群ごとに表11に示す。各評価者群によって不合格率は異なっていた。評価者第1群、第2群、第3群および第4群の評価における歯科技工士学科学生の不合格率は14.8、27.1、35.4および16.7%であり、また歯科技工士専攻科のそれは8.3、8.3、16.7および4.1%であり、歯科技工士学科、歯科技工士専攻科ともに第3群の評価者の不合格率はその他の評価者群の不合格率に比較して高かった。また、いずれの評価者においても歯科技工士学科学生に比較して歯科技工士専攻科学生で不合格率は著しく低かった。

学内成績と不合格者についてみると、歯科技工士学科の学内成績上位(Aランク)の受験者では不合格者が学内成績中位(Bランク)および学内成績下位(Cランク)に比較して著しく少なかった。しかし、学内成績中位(Bランク)および下位(Cランク)の比較では下位の受験者で不合格者数が中位の受験者より少ない傾向にあった。歯科技工士専攻科学生の学内成績上位(Aランク)の受験者では不合格者が0で、学内成績中位(Bランク)および下位(Cランク)に比較して少なかった。

歯科技工士学科学生の学内成績と実技試験成績は必ずしも平行していないことが明らかとなった。

合格基準として、いずれかの試験課題の評価値に25点がある場合を不合格

とする基準を採用した場合の不合格率を表 12 に示す。不合格率は評価者群によって差があったが、不合格率が最も高い試験課題 4 では、不合格率は評価者第 1 群、第 2 群、第 3 群および第 4 群でそれぞれ 27.5、23.3、30.8 および 21.7%であった。

いずれかの試験課題の評価値に 25 点がある場合を不合格とする合格基準を採用した場合の不合格率は比較的高いことが明らかとなった。また、25 点があるため不合格となった受験者と総合計評価値 240 点未満で不合格となった受験者は必ずしも一致しなかった。

## 2) アンケート調査結果について

評価者に対するアンケート調査においては、評価・採点時間について概略的評価には時間を要しなかったが細分化評価では 90%の評価者が、時間がかかったとし(図 10)、評価にあたっては採点基準となるモデルの必要性を述べている。一方、受験者の試験時間に対しては適切な時間であるが 60%以上を占めた(図 11)。また試験問題の内容については、評価者は 75%以上が適切であると答えたが、受験者は 43%であった(図 12)。受験者の多くは歯科技工士学科 2 年生で、資格試験における実技試験をまだ経験していないことから十分把握できず、半数近くがわからないと答えた。今回使用した模擬試験用モデルについては評価者の 50%以上が、受験者の約 80%が不適切と答え(図 13)、その多くは後方からの視認ができない、咬合位の不安定さなどを挙げて

いた。

## D. 考 察

### 1) 本研究実施の背景

歯科技工士資格試験は歯科技工士として必要な知識および技能について行うこととし(歯科技工士法 第 3 章 第 11 条)、試験は厚生労働大臣が毎年少なくとも 1 回行い、試験に関する事務は政令の定めるところにより都道府県知事が行うことができる(歯科技工士法 第 3 章 第 12 条)。現状は歯科技工士学校養成所の設置されている都道府県において毎年 2 月から 3 月の間に行われている。試験内容については歯科技工士法第 2 章第 8 条に基づき 8 科目(歯科理工学、歯の解剖学、顎口腔機能学、関係法規、歯冠修復技工学、有床義歯技工学、矯正歯科技工学、小児歯科技工学)と実地試験(歯科技工実技・実技試験)で、出題基準は歯科技工士養成所教授要綱を参考に作成される。この歯科技工士試験出題基準においては科目ごとに、出題方針を提示し、範囲を定め、各範囲ごとに大項目、小項目に分類している。さらに出題基準は正しく理解され、各都道府県の試験委員によって活用されることによって試験が妥当な範囲と適切な水準で行われることを期待すると結んでいる。しかし、現在の実施されている歯科技工士資格試験の実態は平成 13 年度厚生労働科学研究「今後の歯科技工士に対する養成方策等に関する総合的研究 主任研究者：渡辺嘉一」でも明らかのように、試験の実施時期、学説試験と実地試験の順序、受験会場の条件、受験費用、試験時間、出題形式、問題数など多くの点で地域差があり、バラ

ツキが大きい。

実技試験においては、総義歯排列(2時間30分)、歯冠彫刻(1時間、2歯分)と任意問題(2時間)が行われ、特に任意問題においては8課題のなかから1題のみ出題される。そのために資格試験受験に要する模型の数、その製作に要する時間と労力は、本来行われるべき歯科技工技術教育において大きな支障をきたしているのも事実である。平成13年9月に報告された「歯科技工士の養成の在り方等に関する検討会意見書」にも述べられているように、現状の歯科技工士資格制度における実技試験は、重要な技能評価法であり、今後の技術革新の流れを踏まえて、調査研究を進めるべきであるとし、さらに同作業委員会の意見書においては、歯科技工士教育が歯科技工士試験によって左右されるのは本来の目的から逸脱するものであり、歯科技工士教育に結果的に影響を与えることとなる実技試験を現在の形態で引き続き行う必要性に対して疑問を示す意見もあるが、歯科技工士の業務を踏まえる限り、実技試験の存続を図ることができるよう努力すべきである。ただし、実施にあたっては歯科技工士の国家試験としてふさわしい客観的な評価法が担保された試験とする必要があり、実技試験としての評価を適切に行うための指標開発と具体的な実施方法や客観的な採点基準、出題内容などについて詳細な調査研究を行い、実施手法を確立する必要性を説いている。

国民に良質な歯科医療を安定的に提供するにあたって、質の高い歯科補綴装置を提供することは極めて重要で、その製作を掌る歯科技工士の資質を向上するためには歯

科技工士資格試験の全国統一化を図ることは必然であり、国家試験にふさわしい実施方法や出題内容および評価方法などの条件整備が急務である。また、平成14年度には歯科医師臨床研修に望む歯科医師の技術能力の格差を是正するとともに今後の歯科医師臨床研修の円滑な実施を図り、より質の高い歯科保健・医療を国民に提供する環境の整備と現行の筆頭試験を補完するための技術能力評価試験の導入を目的に「歯科医師国家試験の技能評価等に関する検討会」が設置され、このたびその結果が公表され、技術能力評価試験の出題範囲、試験内容については具体的詳細に示されているものの、その評価方法に関しては「一定レベル以上の経験を有している複数の評価者が行うべきである」ととどまっている。

医療関係職種の中なかでも特に歯科医療においては技能領域の重要性が高く、従来から技能評価としての実技試験の方法や評価法については議論されているが、採点評価結果は評価者の知識や経験による主観的要因が多く含まれることから最近ではコンピュータシステムを応用した評価法も開発、検討されているが、年間2500名の歯科技工士資格試験受験者に対応するには「経済性」「効率性」などの点でまだまだ解決しなければならない問題点も多く、また歯科補綴装置の製作においては単に形態や寸法精度の評価だけではなく、専門家による感性も必要であると考え。また、情報公開の流れを受けて受験者に対しては試験結果の公表も考慮しなければならない。このような観点から複数の専門家による客観的評価法の確立が望ましいと考え、本研究においては、

歯科技工士資格試験の実技評価を行うにあたって客観的評価法の妥当性について多方面から検討した。

## 2) 研究方法について

### ① 受験生および試験用模型・試験課題について

本研究を実施するにあたっては、歯科技工士資格試験の受験を間近に迎える歯科技工士学科 2 学年学生および受験を終えた直後の専攻科(実習科)1 学年学生のうちから、全国歯科技工士教育協議会の協力のもと本研究の主旨、必要性を十分理解できる学生 120 名を抽出した。なお、試験成績と平素の学内成績について検討するために受験した学生についてはあらかじめ平素の成績を A (成績上位)、B (成績中位) および C (成績下位) にランク付けしたが、試験結果や評価結果に影響することを避けるために受験者および評価者には告知していない。今回の資格試験は、「経費面」および「設備面」を考慮して各受験者の所属する歯科技工士専門学校実習室で行い、本研究の主旨を理解している各学校の専任教員が試験説明および試験監督を行った。

今回の試験用模型を策定するにあたっては、従来から行われている歯冠彫刻、総義歯排列(歯肉形成含む)および任意問題の出題内容および方法について再検討し、実施時間や出題傾向などを考慮したうえで、なおかつ実技試験のための準備が歯科技工士学校の教育に負担過剰にならないことを配慮して製作した。すなわち従来の歯冠彫刻は、より臨臨床的配慮に基づき上顎大臼歯全部鑄造冠蟻形成を取り入れ、義歯排列については咬合および審美的な要素を考慮して

上下顎前歯部の人工歯排列と歯肉形成、任意問題として出題されていた試験内容からは部分床義歯を取り上げ、口腔内で喪失率の高い下顎第二小臼歯および第一大臼歯欠損症例における鑄造鉤蟻形成、線鉤屈曲の異なったテクニックの技能を評価することを前提にした。平成 14 年度「厚生労働科学研究 今後の歯科技工士の養成方策等に関する研究」で行ったアンケート調査では従来とおりの総義歯排列や歯冠彫刻の必要性を唱える意見も多かったが、資格試験のための準備や時間的制約などを考慮した場合、今回策定した模型に集約せざるを得ないと考えた。今後は資格試験として統一化を図る場合、他の症例などを組み合わせて、歯科技工士としての基本的技能を諮るにふさわしい試験用模型が試験当日受験者に提供できるよう配すべきであろう。なお、今回の試験用模型製作にあたっては、画一的な模型を多量に提供する必要性から専門の製作者に委託したが、1 回限りの試験、経済性などを考慮して上下顎模型を簡易咬合器に付着したため受験者および評価者から咬合位の不安定、模型後方からの視認不可、模型の扱いにくさなどが指摘された。さらに試験問題として「歯科技工指示書」に基づいて出題したが、受験生の理解度が乏しく、歯科技工士教育における歯科医師との臨床的な情報交換の方法についての必要性を痛感した。

## 3) 評価結果について

### ① 評価法について

評価法として概略的評価法と細分化評価法とを取り上げ、4 校の歯科技工士養成機関において被験者 120 名について模擬的実

技試験を実施し (1) 概略的評価と細分化評価の相関、(2) 概略的評価と細分化評価の構造的解析、(3) 両評価の所要時間について検討した。

受験者 120 名が行った 4 試験課題の答案を評価者 4 群 28 名がそれぞれ 2 つの評価法で評価したので大きなデータ群が得られた。

データの正規性の判定において、総合計評価点は正規確率プロットにより正規性を仮定しうるデータであることが確認された。以後、正規性を前提とした解析を試みた。素データは離散量であるが、多項目、多評価者の評価結果を加算することにより近似的に連続量となり、また大標本数のデータであることから中心極限定理により、正規分布を前提とした解析法によっても定性的な誤りはないと考えられた。ノンパラメトリック手法は、順位列データに対する解析手法であることから 4 段階評価列に対する適用が困難であり、また数量化理論では 4 段階評価を 0、1 評価列にする際に情報量が減るため今回の解析に適用しなかった。

評価者群間の比較は平均値の差の検定により行った。行、列とも要因によって分類に水準化がなされていないため、かつ 1 被験者ごとに全評価者 28 名の評価を比較する必要があることから、従属 2 標本の t 検定を用いた。

概略的評価と細分化評価の相関において、散布図は受験者数が 120 名と大標本であったため、検出力が極めて高かった。概略的評価と細分化評価は、評価時のばらつきの性質は異なるものの回帰直線をもとめてデータを全体としてみると、両者はよく一致していた。

概略的評価と細分化評価の構造的解析においては、受験者 120 名の各試験課題の概略的評価と細分化評価との関係を、概略的評価を従属変数として各々の評価者について重回帰分析を行った。各評価者群が概略的評価にあたって主点をおいた評価項目が明らかとなり、評価者群内のばらつきの原因が明確になった。概略的評価においては評価者第 3 群以外の各評価者群は同様の評価値に収斂するのに対し、細分化評価においては個々の評価項目に関する評価の差が合計することにより拡大したため総合計評価値が発散し、すべての評価者群間の相違が検出された。評価方法と評価一致度および評価項目と評価一致度の間にはかなりの交互作用があると考えられた。

評価の所要時間は、概略的評価においては細分化評価の約 50% の短時間で評価が可能であり、かつ、ばらつきが小さいことが明らかとなった。

以上の結果は、実技試験の評価においては概略的評価法を採用することが妥当であることを示しているものと考えられた。

## ②評価者について

4 群の評価者群、すなわち第 1 群大学教員 (9 名)、第 2 群歯科技工士学校教員 (9 名)、第 3 群臨床歯科医師 (5 名)、第 4 群臨床歯科技工士 (5 名) 計 28 名による概略的評価と細分化評価を実施し、評価者群間の比較を行った。また、6 名の評価者により同一モデルの繰り返し 2 回の概略的評価を実施し、評価者内ばらつきの検討を行った。評価者群間の比較においては、概略的評価では評価者第 3 群以外の各群は同様の評価値に収斂するが、細分化評価では総合計評

価値が発散しすべての評価者群間の相違が検出された。

各試験課題の概略的評価における評価項目ごとの標準回帰係数から、評価者第 1 群および第 2 群では、評価項目 1~5 のすべてを均等に考慮した評価が行われたことを示していた。これに対して評価者第 3 群および第 4 群では、概略的評価がある特定の評価項目に依存する傾向が強く、しかも依存する評価項目が評価者個人によって異なることを示していた。評価者第 1 群および第 2 群では、すべての試験課題にわたって第 3 群、第 4 群に比較してばらつきが少なかった。

測定精度を評価する際用いられる CV 値によって群間比較による評価者群内の評価の一致度を検討した結果は、試験課題 1 (全部鑄造冠蟻形成) に関しては、評価者第 3 群が極めて大きなばらつきを示した。また、試験課題 2 (人工歯排列・歯肉形成) および試験課題 3 (鑄造鉤蟻形成) に関しては評価者第 3 群および第 4 群のばらつきが大きく、第 1 群および第 2 群に対して有意の差が検出された。試験課題 4 (線鉤屈曲) に関してはすべての評価者群間で相違が小さく、よく一致した評価を示した。

試験課題 1 の全部鑄造冠蟻形成においては、評価項目 1 の概形態に対してはいずれの評価者も評価の重要度が高く、評価者群間のばらつきも少ないが、評価項目 5 の表面仕上げについては評価者第 4 群の臨床歯科技工士において評価の重要度が他の評価者群に比べ高い。しかし、いずれの評価者も全部鑄造冠蟻形成の評価にあたっては評価項目の重要度の傾向はほぼ同様であった。

試験課題 2 の人工歯排列においては評価項目 2 の歯軸傾斜に対して評価者群のばらつきは少ないが、評価項目 1 の歯列弓形態では評価者群のばらつきが比較的大きい。また、評価者第 3 群に評価においては特定の評価項目に対する依存が少ない。試験課題 3 の鑄造鉤蟻形成においては評価者第 1 群と第 3 群の歯科医師群で評価項目 3 の形状 (断面形態) に対する評価の重要度が高く、評価項目 4 のレストの状態に対する評価では評価者群のばらつきは少ない。試験課題 4 の線鉤屈曲においては評価者第 3 群では評価項目 3 の屈曲の状態 (キズ)、評価者第 4 群では評価項目 1 の設計線に対する忠実性の重要度がそれぞれ高かったが、概して 4 つの評価者群における各評価項目に対する評価のばらつきはきわめて小さく、線鉤屈曲の評価においては各評価者群の重要度はほぼ同じような評価項目であった。以上のように、全部鑄造冠蟻形成および線鉤屈曲は、各評価者群における評価項目の重要度はほぼ同程度であることから試験課題として評価の安定性が得やすい内容であると考えられる。

評価者個人内のばらつきに関しては、概略的評価を 2 回繰り返した 6 名の評価者中 2 名の評価者において一部の試験課題で繰り返し評価値が一致していたが、その他の評価者では試験課題によって一致あるいは不一致があり、全体として繰り返し測定においては評価値が一致しにくい結果が得られた。これは大標本の高い検出力によるところが大きいと考えられる。しかし、2 回の測定における評価値の平均値の差は最大値で 12.7 であり、評価 1 ランク未満 (25

点未満)であったので判定に大きな影響を与えるようなばらつきとはいえないと考えられる。

以上のごとく、各評価者群は、評価精度、評価所要時間等において特性があった。すなわち、評価者第1群および第2群は、(1) 細分化評価においてすべての試験課題にわたって評価項目1~5のすべてを均等に考慮した評価をおこない、かつ、ばらつきが少ない。(2) 概略的評価における評価精度が他の群より良好である。(3) 評価所要時間は、短時間であり、またばらつきが少ない傾向にある。これに対して評価者第3群および第4群は、(1) 概略的評価においてすべての試験課題にわたって特定の評価項目に依存する傾向にあり、かつ、ばらつきが大きい。(2) 概略的評価における評価精度は試験課題によって異なるが、概してばらつきが大きく他の群より劣る。(3) 評価所要時間は、比較的長時間の傾向がある。

評価者第1群と第3群は、また評価者第2群と第4群は各々の評価項目において同様な評価を行っていた。評価者第1群および第2群のこのような特性は、大学または歯科技工士学校の教育職であり、日頃から評価に熟練していることに起因すると考えられる。特に、第2群は歯科技工士養成を担当しているため実技試験評価においても均衡のとれた良好な評価を行うことができたものと考えられる。評価者第1群と第3群および評価者第2群と第4群の類似した評価特性は、歯科医師、歯科技工士という所属背景が関係しているものと推定される。

実技試験の評価者の選定にあたっては評価者の特性の理解が重要である。また、評

価実施にあたっては、全評価者、特に第3群および第4群の評価者については事前に評価に必要な評価項目について十分な理解とトレーニングを得ることが重要である。

評価者個人内のばらつきに関しては、2回の測定における評価値の平均値の差は最大値で12.7であり、評価1ランク未満で大きな影響を与えるようなばらつきとはいえないと考えられるが、合否基準策定においてはこのばらつきの範囲を十分考慮する必要があると考える。

### ③合格基準について

2つの合格基準、すなわち総合計評価点240点以上合格および各試験課題の評価値に25点があるとき不合格、を模擬試験結果に適用した場合の不合格率について検討した。

合格基準を総合計評価点240点以上としたときの不合格率は、評価者第1群、第2群、第3群および第4群の評価において17.5、23.3、31.7および14.2%で、評価者群で著しい差があった。このことは評価者の評価に対する基準がそれぞれ異なり、評価者の選定および評価レベルの確認を的確に行う必要性を示している。一方、合格基準として、各試験課題の評価値に25点があるときを不合格とする基準を採用した場合の不合格率は、評価者第1群、第2群、第3群および第4群の評価において、それぞれ16.7、15.8、27.1および12.3%であった。各試験課題について120名の受験者における25点評価点の頻度を調べると、25点を獲得した受験者のうち62.8%の受験者が評価者28名中の1名のみによる25点採点であった。このことはこの基準を運用するにあたって、

複数の評価者による再評価が必要であることを示唆している。この基準による不合格者の総合計評価点の中には、240点以上である受験者があったため、25点の評価点があるため不合格となった受験者と総合計評価点240点未満で不合格となった受験者とは必ずしも一致しなかった。

2つの合格基準ともに合格率が比較的低い結果となった。これは受験者の大部分が歯科技工士学科（本科）2年生の第2学期に模擬的実技試験を行ったためで、2年生修了時の学力に十分達していなかったためと考えられる。従って今回の低い合格率をもって実技試験問題のレベルを下げることを検討する必要はないと考えられる。

#### ④実技試験評価と学内成績との関連について

受験者の学内成績は、所属する4校の歯科技工士養成機関が同一基準で定めたものではない、また、学内成績はすべての学科目の総合成績であり、模擬的実技試験評価と比較することが基本的に困難であるが、各養成機関の申告にしたがった歯科技工士学科（本科）学生および歯科技工士専攻科（実習科）学生の、それぞれの成績上位、中位、下位に属する各受験者について概略的評価の総合計評価値と受験者の学内成績の関係を評価者群別ヒストグラムおよび各2群の評価値のt検定により比較した。

歯科技工士学科2年生では学内成績上位の学生の評価値は中位および下位の学生より高点であったが、成績中位と下位の学生の評価値の間には有意の差が認められなかった。歯科技工士専攻科では学内成績上位の学生の評価点が中位、下位より高点とす

る評価者群があったが、3者に評価値の間には有意の差が認められなないとする評価者群もあった。歯科技工士学科2年生と歯科技工士専攻科の受験者全体の評価値は、いずれの評価者群ともに歯科技工士専攻科学生に高い評価値をつけていた。学内成績と模擬的実技試験合格率とは必ずしも平行していなかった。特に、歯科技工士学科の学内成績下位の受験者の不合格率は中位のそれより低かった。

したがって学内成績で資格試験の評価は困難であり、実技試験による評価の重要性が明らかとなった。

#### 4) アンケート調査結果について

評価者に対する評価・採点時間の感想は、概略的評価は時間がかからなかったが、細分化評価には90%以上の評価者は時間がかかったとし、評価結果の分析と一致する。概略的評価においては評価者のこれまでの経験と知識によって各試験課題の作品を総覧的に、主観的に観察することができるために所要時間も短く感じたものとする。また、概略的評価では4つの試験課題が1つの模型に集約されているために、1課題の評価は次の課題の評価に影響を及ぼすと考える。細分化評価ではそれぞれの試験課題において5つの評価項目を設定しているために、丁寧に観察する必要があり、さらに設定した評価項目の重み付けを評価者自身が決定するため、多くの時間を要したと考える。

試験時間に対する受験者の反応は、概して適切であると解答したが、個々の試験課題については短いあるいは長いと判断している。特に人工歯排列では歯肉形成まで完

成するためには2時間30分が必要とする受験生もいる。

試験問題の内容についての多くのコメントは、出題された歯科技工指示書の見方、内容などに関するもので、歯科技工士教育のなかで十分教示する必要性がある。試験課題について、人工歯排列の臼歯部咬合の必要性を訴えるコメントもあり、現在資格試験で総義歯排列が必須化されているなかで、その必要性は認めるものの全国統一された場合の実技試験の実施方法や評価方法あるいは経済的な面などを鑑みた場合、前歯部または片側臼歯部の排列と歯肉形成が妥当と考える。実際の運用面にあたっては、部分的な欠損症例をいくつか準備する必要があると考える。

試験に用いた模型については、評価者および受験者ともに「適切でない」と解答したものが多かったが、ほとんどが咬合位の不安定さ、舌側・後方からの視認性の悪さなどを指摘しているが、今回製作した模型においては咬合関係の必要性、試験当日模型(試験問題)を配布し資格試験のみに使用することなどの経済性から、最も簡易な半透明性のジョイント部を有する咬合器に装着した。また、歯や歯槽部の形態、人工歯の大きさあるいは可撤式歯型など、個々に改善すべき問題もある。

## E. 結 論

歯科技工士資格試験において実技試験は重要な技能評価法として、国家資格の試験としてふさわしい客観的評価基準が担保されたものでなければならない。そこで今回、歯科技工士養成機関の学生 120 名を対象に

模擬的資格試験を実施、所属の異なる評価者 4 群 28 名によって採点評価を行い、評価方法および評価者の違い、合格基準、実技試験評価と学内成績との関連性などについて検討した。その結果、以下のような知見を得た。

1. 試験課題別および評価者群別にみて、概略的評価合計値と細分化評価合計値は相関していることが明らかとなった。
2. 概略的評価および細分化評価の評価値における評価者群間比較によって、概略的評価において臨床系歯科医師の評価者群以外は類似した評価値であることが明らかとなったが、細分化評価においては評価者群間で相違があることが明らかとなった。
3. 概略的評価における各評価者群評価精度は試験課題によって異なっていたが、大学および歯科技工士専門学校の教員群の評価は他の群より良好な精度であることが明らかとなった。
4. 概略的評価と細分化評価の構造解析から各評価者群が概略的評価において細分化評価における各評価項目に依存する度合いに差があることが明らかとなった。
5. 概略的評価における評価者内のばらつきは認められるが、評価に大きな影響を与えるものではないことが明らかとなった。
6. 評価者群間の評価所要時間の比較において大学および歯科技工士学校教員の評価群の所要時間は、臨床系歯科医師および歯科技工士の評価群に比較

して短時間であり、またばらつきが少ない傾向にあった。

7. 概略的評価総合計評価値において、歯科技工士学科学生の学内成績上位者の評価値は中位者および下位者より高点で有意の差が認められたが、歯科技工士専攻科学生の学内成績と試験評価値との間には有意の差は認められなかった。また、いずれの評価群においても受験者全体の評価値は歯科技工士専攻科学生のほうが歯科技工士学科学生より高い評価値を示した。

8. 合格基準を総合計点数の 60%とした場合、不合格率は評価者群によって異なり、臨床系歯科医師の不合格率はその他の評価者群に比較して高かった。また、いずれの評価者群においても歯科技工士専攻科学生の不合格率は歯科技工士学科学生より著しく低かった。さらに、いずれかの試験課題に評価値が最低点を含む場合を不合格とする合格基準を設定した場合、不合格率は比較的高くなることが明らかとなった。

以上の結果から、歯科技工士資格試験の実技試験評価にあたっては、評価精度ならびに時間的要素から複数の評価者による段階的評価法による概略的評価が望ましく、評価者は十分なトレーニングと評価項目に対する理解と評価レベルの確認を行うことが重要であり、合格基準の設定にあたっては十分な配慮が必要で、不合格者に対しては複数の評価者による再評価の必要性が明

らかとなった。さらに学内成績で資格試験の実技評価は困難であり、実地試験による技術評価の重要性が示唆された。

#### F. 健康危険情報

なし

#### G. 研究発表

なし

#### H. 知的所有権の出願・登録状況

なし

#### I. 参考文献

- 1) 厚生労働省：歯科技工士の養成の在り方等に関する検討会意見書，2001.
- 2) 厚生科学研究 医療技術評価総合研究事業 今後の歯科技工士に対する養成方策等に関する総合的研究 研究報告書，渡辺嘉一，2003.

#### J. 研究協力者

- 渡辺嘉一 日本歯科大学歯学部  
新井 泉 東京医科歯科大学学生体材料工学研究所  
小黑 章 明倫短期大学  
尾崎順男 日本歯科大学附属歯科専門学校  
杉上圭三 大阪歯科大学歯科技工士専門学校

表1 試験内容と試験時間

課 題	時 間
部分床義歯の人工歯排列および歯肉形成 $\begin{array}{r l} 321 & 123 \\ \hline 321 & 123 \end{array}$	1時間30分(90分)
15分間休憩	
全部鑄造冠蟬形成 $\frac{ 6}{7 }$ 鑄造鉤蟬形成 $\frac{7 }{4 }$ 線鉤屈曲	2時間(120分)

表2 試験評価者の区分

所 属	人 数
第1群 大学教員(補綴系教授・助教授・講師)	9名
第2群 歯科技工士専門学校教員(教員歴15年以上)	9名
第3群 臨床歯科医師(開業歯科医師 48歳以上)	5名
第4群 臨床歯科技工士(開設歯科技工士 40歳以上)	5名

表3 細分化評価の各評価項目における  
着眼点

試験課題1 : 全部鑄造冠蟬形成

評価項目
1. 概形態 ① 咬合面形態(咬頭の位置・大きさ) ② 頬・舌面形態(咬頭の位置・豊隆) ③ 近・遠心形態(辺縁隆線の高さ) ④ 完成度
2. 咬合状態 ① オーバージェット・オーバーバイト ② 咬合接触(有無・位置)
3. 接触点 ① 有無・位置・大きさ
4. マージン辺縁部 ① 適合性(オーバー・アンダーマージン) ② 豊隆・移行状態 ③ 完成度
5. 表面仕上げ ① 仕上げ研磨(滑沢度) ② 完成度

試験課題2 : 人工歯排列

評価項目
1. 歯列弓形態 ① 人工歯の部位・位置・方向 ② 歯列弓の概形態 ③ 完成度
2. 歯軸傾斜 ① 歯種による近遠心・頬舌傾斜方向
3. 垂直・水平被蓋関係 ① 被蓋関係の大きさ
4. 歯頸線の位置 ① 高さ ② 連続性
5. 歯肉形成 ① 豊隆度 ② 辺縁形態・位置 ③ 仕上げ状態(研磨) ④ 完成度

試験課題3 : 鑄造鉤蟬形成

評価項目
1. 設計線に対する忠実性 ① 設計に対する理解度 ② 設計線への忠実性 ③ ワックスの取り扱い
2. 概形態 ① 鉤先・鉤腕・鉤体・鉤脚の形態 ② 完成度
3. 形状 ① 断面形態(厚さ・幅)
4. レストの状態 ① 形態 ② 咬合接触状態 ③ 完成度
5. 表面仕上げ ① 仕上げ研磨(滑沢度) ② 完成度

試験課題4 : 線鉤屈曲

評価項目
1. 設計線に対する忠実性 ① 設計に対する理解度 ② 設計線への忠実性
2. 歯面への適合性 ① 歯面への接触状態
3. 屈曲の状態 ① スムーズな屈曲 ① プライヤーによるキズ
4. レストの状態 ① 形態 ② 咬合接触状態 ③ 完成度
5. 鉤先の処理 ① 鉤先の位置 ② 研磨状態 ③ 完成度