

WHOのチェックリストを用いた日本版

「手術安全簡易評価システム」の開発と適応に関する研究  
- 外科手術におけるノンテクニカルスキルの評価の現状についての文献的検索 -

研究代表者 兼見 敏浩 三重大学医学部附属病院 教授

## はじめに

航空、原子力工学などのハイリスク産業分野において、ひとたびトラブルが発生すると、人的・物的・そして環境への甚大な被害が予想される。実際に東日本大震災時の東京電力福島第一原子力発電所事故による被害は甚大であり、未だに収拾の目処が立っていない。したがって、これらのハイリスク産業においては、各組織はリスクを回避するために設備・システムといったハード面の技術開発と従業者の能力開発に多大な投資を行ってきた。その結果、特に近年はハード面での進歩が著しく、多重的防護システムが構築されつつあることから、かつてのように“機械・設備の不具合による重大な事故”は減少傾向にあり、トラブル発生時には人的な要因が増加しつつあると考えられる。従業者の能力開発もハード面の技術開発と同等に積極的に行われてきたが、その対象は専門技術すなわち、テクニカルスキルが中心であった。しかし、事故の防止にはテクニカルスキルのみならず、状況認識・意思決定・コミュニケーションとチームワーク・リーダーシップといったノンテクニカルスキルも重要であることをいち早く認識した航空業界では、「Crew Resource Management (CRM)」という研修課程を通じて重要なノンテクニカルスキルを経験的に特定し、その習得のための訓練を行ってきた[1,2]。医療分野の有害事象分析においても、個人のテクニカルスキルが高い水準でプラトーに達した状態では、未熟なノンテクニカルスキルが複雑化する組織構造における有害事象発生の重要な要因となることが明らかとされている。特にこの20年間は、安全で良質な医療の提供におけるノンテクニカルスキルの重要性に対する認識が高まってきた[3-7]。

## ノンテクニカルスキルの重要性

医療分野では、まず、航空機のcockpitに近い環境である麻酔科領域でノンテクニカルスキルを評価するための行動マーカーシステムが“Anaesthetists' Non-Technical Skills: ANTS”という行動マーカーシステムが開発されていた[8]。同様に、外科医のノンテクニカルスキルを評価するための各種ツールも開発されている。このような関心の変化が起こった理由は、術中有害事象の分析において、その要因が技術的な専門知識の欠如だけでなく、ノンテクニカルスキルの欠如にもあることが指摘されたためである[9-13]。わが国の日本医療機能評価機構による医療事故情報収集等事業においても、医療事故の原因として、技術が未熟だった、知識がなかったといったテクニカルスキルが推定されたものは少数であり、多くは確認を怠った、観察を怠ったなどのノンテクニカルスキルが原因であった[14]。

発生要因	2012年1月～12月	
	件数	%
当事者の行動に関わる要因	44,625	56.4
確認を怠った	20,237	25.6
観察を怠った	7,110	9.0
報告が遅れた(怠った)	891	1.1
記録などに不備があった	894	1.1
連携ができていなかった	5,367	6.8
患者への説明が不十分であった(怠った)	4,047	5.1
判断を誤った	6,079	7.7
ヒューマンファクター	17,744	22.5
知識が不足していた	2,901	3.7
技術・手技が未熟だった	1,886	2.4
勤務状況が繁忙だった	6,838	8.7
通常とは異なる身体的条件下にあった	416	0.5
通常とは異なる心理的条件下にあった	2,537	3.2
その他	3,166	4.0
環境・設備機器	7,612	9.7
コンピュータシステム	1,113	1.4
医薬品	1,460	1.9
医療機器	596	0.8
施設・設備	503	0.6
諸物品	557	0.7
患者側	2,471	3.1
その他	912	1.2
その他	8,917	11.3
教育・訓練	2,748	3.5
仕組み	585	0.7
ルールの不備	1,249	1.6
その他	4,335	5.5
合計	78,898	100.0

**図1 医療事故の発生要因**

2012年に日本医療機能評価機構に対して報告義務対象医療機関(273)および参加登録申請医療機関(653)から報告された医療事故の発生要因を示す。発生要因として“知識が不足していた”、“技術・手技が未熟だった”といったテクニカルスキルに属するものは、それぞれ、3.7%、2.4%にすぎず、多くは、“確認を怠った”(25.6%)、“観察を怠った”(9.0%)などのノンテクニカルスキルに属するものであった。参考文献[14]を一部改変。

2012年に日本医療機能評価機構に対して報告義務対象医療機関および参加登録申請医療機関から報告された医療事故の発生要因一覧を示す(図1)。発生要因の多くは知識が不足していた”、“技術・手技が未熟だった”といったテクニカルスキルに属するものではなく、確認を怠った”、“観察を怠った”などのノンテクニカルスキルに属するものであった。

## 外科領域におけるノンテクニカルスキルの評価システム

過去10年間に、外科領域におけるノンテクニカルスキルの評価を目的とする2つの主要な能力分類基準、すなわちNOTSS[15,16]とNOTECHS[17]が開発された。いずれの評価システムも、手術室内で発生したノンテクニカルスキルに関連する過誤に関する研究や、航空産業や原子力産業といった高リスク産業で用いられている評価ツールによって得られた、ノンテクニカルスキルの重要性を強調する教訓に基づいている。さらに、外科領域におけるチームワークの評価を目的としたツールOTAS[18]や、外科以外の分野におけるNTS評価ツールも開発されている[2]。ここではNOTSS、NOTECHS、OTASの特性を比較検討してみる。

### NOTSS (Non Technical Skills for Surgeons)

#### 開発の経緯

NOTSS (Non Technical Skills for Surgeons) は、外科医の手術中におけるノンテクニカルスキルに焦点を当てて開発された評価スケールである[16]。このスケールは2006年に開発され、さまざまな外科専門分野の上級外科医27名を対象として、クリティカル・インシデント法という聴取法を用いて行われたインタビューの結果に基づいて策定されたものである。このインタビューでは、外科医が術中の出来事や行動を思い出すことができるように、いくつかの特殊な質問方法が用いられた。このインタビューは、認知能力を明らかにすることだけでなく、手術室内でクリティカル・インシデント(重要な出来事)が発生した場合に各外科医が使用している対人関係スキルに注意を向けることを目的としている。インタビュー終了後、3組の心理学者がそれぞれ独立にインタビュー内容を分析した。分析は、グラウンデッド・セオリーに基づく逐次処理符号化法を用いて、評価者間信頼性が許容可能な水準に達するまで行われた。この分析は、麻酔領域と欧州の民間航空会社で既に用いられていた系統的プログラムデザインとともに、NOTSSの開発に貢献した[19,20]。さらに、手術室内の過誤に関する文献考察、手術室で働く人員の態度調査、死亡報告の分析、及び手術室内における観察結果も、本評価システムの開発に寄与した。上級外科医に対するインタビューの結果から、ノンテクニカルスキルの主要なカテゴリーとして、「situation awareness: 状況認識」、 「decision-making: 意思決定」、 「task management: タスク管理」、 「leadership: リーダーシップ」、 「communication and teamwork: コミュニケーションとチームワーク」の5つが特定された。タスク管理以外の4つのカテゴリーに3個の特異的なスキル要素が存在し、スキル要素の総数は12個となる(図2)。観察可能な行動に基づき、各要素を1点(不良)~4点(良好)の4段階で評価する。なお、当該行動が特定のケース又はシナリオにおいて必要とされない場合は、「NO」と記入する。各要素の評価スコアを合計して総客観スコアを求め、外科医のノンテクニカルスキルの指標とする(図3)。

また、Yuleらは16名の上級外科医に密着して調査を行った結果、NOTSSツール使用時の評価者の参考となる、良好なパフォーマンス行動と好ましくないパフォーマンス行動の例を報告した[16]。例えば、状況認識カテゴリーの要素の一つに情報収集があるが、この要素に関する良好なパフォーマンスを示す観察可能な行動マーカーの一例として、「すべての重要な検査結果の検討が行われていることを確認する」という行動が挙げられる。反対に、「手術室に遅れて到着する」、あるいは「手術室に来よう繰り返し呼ばれる」という行動は、好ましくない行動とみなされる(図4～)。

カテゴリー Category	要素 Element
状況認識 Situation awareness	<ul style="list-style-type: none"> <li>□ 情報を集める: Gathering information</li> <li>□ 情報を理解する: Understanding information</li> <li>□ 先を見通し、行動する: Projecting and anticipating future state</li> </ul>
意思決定 Decision-making	<ul style="list-style-type: none"> <li>□ 選択肢を検討する: Considering options</li> <li>□ 選択を行い、チームに伝える: Selecting and communicating option</li> <li>□ 選択を実行し、経過を確認する: Implementing and reviewing decisions</li> </ul>
コミュニケーションとチームワーク Communication and teamwork	<ul style="list-style-type: none"> <li>□ メンバー間で情報を交換する: Exchanging information</li> <li>□ 相互的な理解をつくりあげる: Establishing a shared understanding</li> <li>□ チームの活動を調整する: Coordinating team activities</li> </ul>
リーダーシップ Leadership	<ul style="list-style-type: none"> <li>□ パフォーマンスの水準を設定しそれを維持する: Setting and maintaining standards</li> <li>□ メンバーをサポートする: Supporting others</li> <li>□ チームのプレッシャーに対処する: Coping with pressure</li> </ul>

**図2 NOTSSのカテゴリーと要素**  
(参考文献[16]円谷ら訳)

表記	内容
4 – Good 良い	手術の遂行は一環して高い標準を維持、患者の安全も促進、良い見本足りうる
3 – Acceptable 普通	標準的で満足行くレベルだが、向上の余地あり
2 – Marginal もう一歩	心配な要素あり、かなり向上する必要あり
1 – Poor 悪い	患者の安全を脅かす、あるいは潜在的に危険な要素あり、大いに改善を要する
N/A –Not Applicable 該当せず	このケースでは”Skill”は不要

**図3 NOTSSの評価スケール**  
(参考文献[16]円谷ら訳)

## 図4 NOTSSにおける良い行動と悪い行動の例

(参考文献[16]円谷ら訳)

### カテゴリー：状況把握

カテゴリー：状況認識		
周囲のデータ(患者、チーム、時間、モニター、機器)を総合して、手術室の状況を刻々と把握できるようにして、かつ継続；それらが何を意味するかを理解、また次に何が起きるか予測		
要素	良い行動	悪い行動
情報を集める： 手術所見、室環境、機器、人々から手術室での情報を求める	検査結果・同意書等カルテの術前チェック 全ての所見(例:画像)レビュー、利用可能 麻酔プランについて、麻酔科と連絡し合う 開始前に、手術に最適な状態にする(例:手術台・ライト・AV機器； S A N) 解剖と病理所見を、明確に見極める 出血をモニター、麻酔医に最新状況を聞く	手術室に遅れるか何度もコールされる 検査結果等をぎりぎりに確認、またはしない 他メンバーの見解を考慮しない 麻酔医に状況を聞かない チームで収集した情報をレビューしない 手術開始前にみたことを理由に、術中にカルテの情報を読むよう頼まない
情報を理解する： 収集情報を解釈しイメージを更新、予定と現状の食い違いを既知と比べて見分ける	術前と手術所見を総合して行動する CTを見て、関連する領域を指摘する 情報の意義について熟考して、話し合う ( S A N )	重要な結果を見逃す、または無視する 明らかなサインを見逃す(例:CT所見) 理解不足を証明する質問をする “青写真”に合わない結果は無視
先を見通し、行動する： 考えられる行為、介入、非介入による転帰を予測	手術や麻酔困難による遅れを考慮し計画 手術の後半で“ ”が必要かもと伝える 臨時プランの根拠を示す(プランB、例:NSに必要な機器が使用可能かを確認) 予期される臨床イベントの文献をあげる	失敗しそうな操作にも、過信している 潜在する問題について、議論しない 予測出血量になってから、麻酔医に伝える 予期される問題が起きてから、対応する 経験以上の難易度の手術をする

S: 外科医、A:麻酔医、N:看護師 (Option)

### カテゴリー：意思決定

カテゴリー：意思決定		
状況を診断して、適切な手順を選択するための判断に到るスキル		
要素	良い行動	悪い行動
選択肢を検討する： 他の手段の可能性・問題解決手順の提起、潜在オプションの危険要素を評価して、リスク便益を理解	問題点を把握、はっきり言う 他メンバーとのオプション・pros and cons'に関するバランスある 討議を開始 メンバーの意見を求める ガイドラインについて討議する	オプションについて討議しない メンバーの意見を求めない ガイドラインを無視する
選択を行い、チームに伝える： 問題解決法を選択し、他の関連メンバーに周知	決断に至り、はっきりと伝える “プランB”の準備をして伝える 何故臨時対応策を選んだかを説明する	手術プランをチームに伝えない プランへの疑問に怒るもしくは無反応 他の治療オプションの議論を封鎖 思い通りにならないと、手術をやめる 合併症につながる不適切な手技を選ぶ
選択を実行し、経過を確認する： 選択手順に着手、患者の状態変化に照らし妥当性を随時見直し。状況変化に対応し必要に応じ柔軟にプランを変更、目標へ	決定事項を遂行する 患者の状態変化に照らして、または問題が生じた場合には、プランを見直す “プランA”がうまくいかないことを悟り、“プランB”に切り替える 進行中にチーム更新、必要時助け要請	決定事項を遂行しない 繰り返し同じミスをする 行為の影響を再考しない 悪い結果が予測されたり、別法により良い証拠があるのに、“プランA”に固執 時間制限を気にして、急いだりあわてる

## カテゴリー：コミュニケーションとチームワーク

カテゴリー：コミュニケーションと チームワーク “チーム”として働くためのSkillで、チーム内での適切な状況認識の共有、効果的な職務遂行を確保		
要素	良い行動	悪い行動
メンバー間で情報を交換する： 知識と情報をタイムリーにやり取りし、 メンバー間の相互的な理解に寄与	心配な点について、メンバーに耳を傾ける 手術の進行状況について話す S (手技、出血、予測AE、他) A・N (患者状態、使用薬剤、他) 手術が計画通りではないことを伝える	心配な点について、協議できない 1人で問題を解決しようとする メンバーの意見を聞かない 助手に助けを求めるが、何をすべきか明確にせず
相互理解をつくりあげる： チームが手術遂行上必要な関連情 報を把握し、かつ、それをメンバーが 理解し、症例の全体像が適切に共有 されていることを保証	手術開始前に、要点[状況]説明をして、目的とゴールを明確に する 手術計画が理解されている事を確認 全メンバーから、追加情報を促す 関連者が、決定に満足であることを確認 手術直後に総括して、何がうまく行き、どんな問題があったか振 り返り ( S A N )	チーム員にはっきりと手術プランを言えない 全体議論や、進捗確認の時間をとらない 慣れないメンバーと事前症例検討をしない 麻酔医に適宜手技の進行・リスクを伝えない 手術プランに異論があると、不愉快にみえる、手術後に問題 点や成功点について議論しない
チームの活動を調整： 他のメンバーと協働して、認知 <sup>†</sup> ・肉 体活動を、同時に協力的に遂行	他のメンバーも手術開始の準備済みであることを確認 麻酔医やNSの求めに応じ手術を止める タイムリーに業務を整理、“チームが効率よく機能”を確認	麻酔医に手術を開始してよいか聞かない 機器が準備できていなくても、手術を始める

† 認知: 感覚・知覚・記憶・認知

## カテゴリー：リーダーシップ

カテゴリー：リーダーシップ チームをリードして方針を示し、高水準の実地医療とケアを実演し、かつメンバー個々のニーズも思いやる		
要素	良い行動	悪い行動
パフォーマンスの水準を設定しそ れを維持する： 適切な手術の基本遵守・臨床試験 の規準 ( G C P ) ・手術室のプロ トコルを守り、安全と質を維持	新しい・馴染みの薄いメンバーに自己紹介 手術室のプロトコルに、明確に従う 全チームに、標準に従うよう要請 ( 例：清潔野 )	標準に従わない ( 例：器具が汚染・不適當であっても 続ける ) 手術室のプロトコルを破る 患者に対して、無礼な態度をとる
メンバーをサポートする： メンバーに認知的・感情的な援助 を提供。個々の能力を判断して リーダーシップの型を合わせる	研修医のニーズに応じ行動を修正 メンバーに建設的な批評を与える 業務委任が適切なことを保証 メンバーと親密な関係を確立 良くなされた職務は功績とする	メンバーのニーズを見分けられない 技術的に‘トンネル視野’ 進入にこだわる 他のチーム員との、敵対を顔に出す ( 例：看護師に対 する皮肉 ) 良くなされた職務を無視
チームのプレッシャーに対処す る：プレッシャー下でも穏やかな 挙動を保ち、チームに切迫状況の 制御を強調。 状況に応じ、他メンバーを害さな い範囲で強制的な作法をとる	プレッシャー下でも、穏やかに 切迫状況を強調 ( 例：時には大声で ) 緊急・重大局面で患者への責任を持つ プレッシャー下でも適切な決断 目標達成のためには、業務を委任 緊急時一貫してチームを統率	臨床上の問題に関する心配を抑制 ‘フリーズ’ してプレッシャー下での決断能力のなさを 露呈 技術的難問で全集中が必要な局面で、リーダーシップ を委任できない 自分以外の失敗を責め、責任をとらない かんしゃくを起こす

## NOTTS 評価システムの検証

NOTSS の評価ツールとしての信頼性と感度については、11 の手術室シナリオのシミュレーションに基づく評価が行われている[21]。各シナリオはビデオ録画され、模擬患者、担当外科医、麻酔科医、看護師がそれぞれ主要な役割を演じた。これらのシナリオは、ノンテクニカルスキルに関する研修を受けた外科医 2 名、心理学者 2 名、及び麻酔科医 1 名によって設計され、多様な臨床シナリオにおける NOTSS の信頼性を評価するために、良・不良を含むさまざまな種類のノンテクニカルスキルが提示された。これらの参加者の大多数(84%)は、受講者のパフォーマンス評価に関する何らかの経験を有していた。本研究に登録された上級外科医は、ヒューマンファクターと NOTSS による行動の観察及び評価に関する簡単な講習(2.5 時間)を受けた。

この評価スケールはタスク分析、インタビュー、文献考察に基づいて作成されたものであるため、その内容妥当性は担保されていると考えられる。また、この評価スケールは良好な行動パフォーマンスと好ましくない行動パフォーマンスを区別することができるため、その構成概念妥当性も良好であると考えられる[2]。

一連の「参照評価」(行動及びノンテクニカルスキル・テクニカルスキルの評価における 10 年程度の経験を有するシナリオ設計者による評価)を用いた感度評価として、研究参加者による良好な行動パフォーマンスと好ましくない行動パフォーマンスの評価が、専門家による評価とどの程度類似しているかが検討された。いずれの論文でも、許容可能なレベルの感度は報告されなかった。これらの報告には、すべての参加者による評価と専門家による評価との差が 1 点未満であった報告も含まれる。感度が最も高かったカテゴリーは「タスク管理」、最も低かったカテゴリーは「状況認識」であったが、「状況認識」は臨床的に最も重要なカテゴリーの一つであるといえる。また、評価スケールを 4 段階から 2 段階に縮小し、カテゴリーを「許容可能」又は「許容不能」のいずれかで評価すると、すべての NOTSS カテゴリーにおいて、63~82%の評価者による評価が専門家による評価と一致した。

カテゴリー評価と要素評価の間の信頼性は高く、さまざまなスキル要素とそれらの行動マーカが、実際にそれらのカテゴリーとリンクしていることが示された。評価者間信頼性は、2 つのカテゴリー(「リーダーシップ」と「コミュニケーション」及び「チームワーク」)間でのみ許容可能なレベルであり、追加の信頼性指標として級内相関を使用すると、いずれのカテゴリーでも許容可能なレベルの評価者間信頼性は認められなかった。さらに、評価者の外科専門分野が多岐にわたっていたことから、分野ごとの信頼性についても個別に評価を行った。その結果、本検討では、整形外科医よりも一般外科医のスコアに大きなばらつきが認められた。また、本評価システムの最終開発段階で「タスク管理」カテゴリーが削除され、関連する行動項目が他のカテゴリーに組み込まれた。これは当該カテゴリーの評価者間信頼性が低かったことと、タスク管理は手術中よりも主に術前計画時に割り当てられたという外科医からのフィードバックに基づく判断であった。これにより評価ツールがよりシンプルとなった。

Yule らによる追加研究では、初心者(2.5 時間の NOTSS 研修を受けただけの上級外科医)と専門家(行動評価及び技術的・非技術的能力の評価における 10 年以上の経験を有する手術チームのメンバー)の評価者による NOTSS 評価スコアが比較された[22]。その結果、初心者評価者の約半数の評価結果が専門家によるビデオ録画された臨床シナリオシミュレーションと一致せず、一致しなかつ

たケースの大部分では、初心者によるスコアは専門家によるスコアよりも低かった。このような専門家評価者と初心者評価者の間でのスコアの不一致は、判定が困難なノンテクニカルスキルパフォーマンスの範囲で比較的多くみられ、明らかに許容不能かつ不安全的な行動については、スコアの相関性はより高かった。これは、初心者評価者グループにおける行動評価の経験不足に起因していると考えられる。また、不一致の多くは、判定対象のビデオに明らかに優れた行動や極端にネガティブな行動パフォーマンスが描写されていない、中間領域の行動カテゴリーで発生していた。初心者評価者は、あるタスクが適切に実行されなかったが、全体的な患者ケアへの影響がなかった場合に、どのように行動カテゴリーを判定すべきかに確信が持てなかった。

その後、Crossleyらは、事前録画した映像ではなく、実際の手術の観察でNOTSSシステムを検証した[23]。麻酔科医、手術室看護師など簡易なトレーニングを受けた初心者評価者とNOTSSトレーニングを受けた専門家評価者が404手術事例を実際に見て行った評価結果を検証したものであるが、その結果、NOTSSシステムは実際の手術現場においても適応可能であること、初心者評価者であっても評価可能であったとしている。さらに、自国の実情に合わせた改良版NOTSSの考案もなされている[24]。

以上よりNOTSS評価システムは有用であり、実際の手術においても外科医の個人としてのノンテクニカルスキルを評価することが可能であると考えられるが、独立した評価者が必要であり、専門家評価者を必ずしも求めるわけではないが、初心者が評価者をする場合には、評価方法の理解や練習のために最低限の時間が必要である。しかしながら、NOTSSは評価をすることだけが目的でなく手術にかかるチーム全体のスキルアップのためのツールとして活用可能である[25]。

## NOTECHS (Non Technical skills)

### スケール開発の経緯

NOTECHSは、航空業界のCRMへの応用を目的として開発されたものである[26,27,28]。この分野での良好な結果を受けて、NOTECHSは麻酔分野[19]を皮切りに様々な分野に応用され、現在では外科分野でも利用されている。

NOTECHSはノンテクニカルスキルを「リーダーシップ及び管理」、「チームワーク及び協力」、「問題解決及び意思決定」、「状況認識」の4つのカテゴリーに分類し、各カテゴリーを1点（非常に不良）から5点（非常に良好）の5段階で評価するシステムである。経験的なエビデンスから、NOTECHSが高い信頼性をもってCRMに応用可能であることが示唆されたことから、外科領域への応用が促された[29,30]。

Sevadalisらは、NOTECHSスケールを手術室での使用により適したものに改良した[17]。彼らは5番目のカテゴリーとして、「コミュニケーション及び相互交流」を追加した（図5）。この変更は、手術室環境とコックピットとの違いに関する理解に基づいて行われた。手術室は看護師、麻酔医、外科医という、多様な訓練背景と職業文化を持つ3種類の専門家から構成されるため、多様な環境が形成される。このことは一般的に、結束力の低下と技術的多様性の増加をもたらす。また、手術室チームの構成員同士は互いに初対面である場合もあり、交代輪番制のため、過去に一緒に仕事を



したことがないという状況で、その場限りのチームとして活動する場合が多い。なお、「コミュニケーション及び相互交流」というカテゴリーの追加は、Healeyらによって確立された、実際の手術室におけるチームワークの理論的枠組みに基づいている[31]。この枠組みの中では、主要なコミュニケーション関連タスクの遂行と、コミュニケーション関連行動の十分なパフォーマンスが、リアルタイムな活動におけるチームワークの総合評価において重要な意味をもつ。改良版のスケールは、1点（行われていない）から6点（非常に良く行われた）の6段階で評価される。

<b>Revised NOTECHS scale for the surgical group</b>	
<b>Subscales</b>	<b>Items</b>
Communication and Interaction	A1. Instructions to assistant clear and polite A2. Waited for acknowledgement from assistant A3. Instructions to scrub nurse clear and polite A4. Waited for acknowledgement from scrub nurse
Situation Awareness and Vigilance	B1. Monitored patient parameters throughout procedure B2. Awareness of anesthetist B3. Actively initiates communication with anesthetist during crisis
Cooperation and Team Skills	C1. Maintains positive rapport with whole team C2. Open to opinions from other team members C3. Acknowledges contribution from other team members C4. Supportive of other team members C5. Conflict handling (concentrating on what is right rather than who is right)
Leadership and Managerial Skills	D1. Adherence to best-practice during procedure (eg<comma> does not permit corner cutting) D2. Time management (eg<comma> not being too slow or rushing other team members) D3. Resource utilization (eg<comma> appropriate task load distribution and delegation of responsibilities) D4. Debriefing the team (eg<comma> provides details and feedback to the team about procedure) D5. Authority and assertiveness
Decision Making	E1. Prompt identification of the problem E2. Informed team members promptly and clearly E3. Outlines strategy and institutes a plan (eg<comma> asks scrub nurse for suction<comma> instruments<comma> suture material) E4. Anticipates potential problems and prepares contingency plan (eg<comma> ask anesthetist to order blood<comma> call for help) E5. Option generation (eg<comma> takes help from others<comma> seeks team's opinion)

**図4 改訂版 NOTECHS 評価スケール（外科手術用）**

Situation Awareness and Vigilance（状況認識と警戒）、Cooperation and Team Skills（協力とチームワーク）、Leadership and Managerial Skills（リーダーシップと管理技術）、Decision Making（意思決定）、に Communication and Interaction（コミュニケーションと相互交流）が追加された[文献 17]。

## 評価システムの検証

NOTECHS スケールは、航空分野における妥当性が多くの研究で実証されているため、そのヒューマンファクター測定ツールとしての内容妥当性は担保されていると考えられる。外科領域では、この評価尺度によって良好な行動パフォーマンスと好ましくない行動パフォーマンスが区別できることが示されているため、その構成概念妥当性も良好であると考えられる[17]。

改良版 NOTECHS については、手術室における危機シナリオのシミュレーションに基づく検討が行われている[17]。本尺度は優れた信頼性を持つことが示されており、クロンバックの係数は5つの下位尺度のすべてで $>0.7$ であった。また、指導者と研修受講者との間で評価結果に大きな差がみられなかったため、自己評価ツールとしても有用であると考えられる。ノンテクニカルスキル研修の前後にかかわらず、複数回の評価間での信頼性も良好であった。さらに、連続して評価を行った場合の信頼性も良好であり、看護師、麻酔医、外科医のいずれが評価者となった場合も有効性に差は認められなかった。

NOTECHS は既に臨床応用されている。手術室内のチームワークのパフォーマンス評価に対応したバージョンである Oxford NOTECHS システムは、オリジナル版と同じ4つのカテゴリーからなり、いくつかの興味深い結果が報告されている。Mishra ら[32]による検討では、手術チームが26件の腹腔鏡下胆嚢切除術と22件の頸動脈内膜摘除術を施行し、術中のノンテクニカルスキルとテクニカルスキルが比較された。技術的エラーの定量法として、OCHRA (Observation Clinical Human Reliability Assessment: 観察臨床ヒト信頼性評価) ツールが用いられた。外科研修医1名と人的要因評価の専門家1名の計2名が観察者となり、手術中の外科医、看護師、麻酔医の行動を観察して個別にスコア化した。外科医、麻酔医、看護師の平均スコアは、16点満点中、それぞれ13.3点、11.4点、10.8点であった。専門家評価者による評価と初心者評価者による評価との間の信頼性は良好であり、クロンバックの係数は0.88であった。術野の外で起こったエラー、すなわち重要な機器の作動不良や器具の設置不備などについては、「Non-Operative Procedural Errors (NOPE)」として記録された。安全意識調査結果、手術時間、患者の入院期間、及びその他の転帰も記録された。研究参加者は、異なる2つの手術の間にノンテクニカルスキル研修を受けた。本検討の結果から、外科医についてはヒューマンファクタースキル研修を受けた後もノンテクニカルスキルスコアに明らかな変化は認められなかったが、これはベースラインにおけるパフォーマンスが高かったためであると考えられた。技術的パフォーマンスとNOPE数はノンテクニカルスキル研修後に低下した。テクニカルスキルエラーとノンテクニカルスキルとの間の全体的な相関性は弱かったが、外科医の状況認識度と技術的エラー発生率との間には強い相関が認められた。手術時間と患者の術後入院期間には、ノンテクニカルスキル研修後も変化は認められなかった。さらに Oxford NOTECHS システムを外科医、麻酔科医、看護師のサブチームに分けてそれぞれに行動評価の詳細を記載したシステムの有効性も報告されている[33]。

以上の結果から、ヒューマンファクター研修はノンテクニカルスキルを向上させ、手術室でのエラー発生率を低下させたが、これはチームワークとコミュニケーションの改善によるものであると考えられた。しかし、本検討は対象例数が少なかったことから、これらの改善は患者入院期間や手術時間の明らかな短縮に結びつかなかった。今後はより大規模な検討において、これらの結果を検

証する必要がある。

NOTECHS 評価スケールは、広範なノンテクニカルスキルの評価に対応している。各カテゴリーに優れた行動記述子が設定されているため、使用にあたって必要とされる研修量は比較的少ない。また、このスケールは研修生と専門家評価者との間でも高い信頼性を示すため、他の評価スケールと比較して、大幅に低いコストで自己評価に応用することができる。さらに、予備的検討で示されている通り、NOTECHS はチーム及び個人の評価やフィードバックにも利用することができ、リアルタイムで使用することができる。

## OTAS (Observational Teamwork Assessment for Surgery)

### 開発の経緯

OTAS は、患者転帰に影響を及ぼす因子を幅広く評価する必要性に基づき、2006 年に開発されたスケールである [18]。またこのスケールでは、個々のチームワークスキルの範疇を超えた、手術室におけるチームワークの評価に重点が置かれている。航空業界で確立され、チーム理論に関する文献にも多く報告されている、チームパフォーマンスに関する基礎的な入出力モデルを応用して、外科チームワークモデルが作成された。

OTAS には、タスクチェックリストとチーム行動評価の 2 つの評価要素がある。タスクチェックリストは、診療ガイドライン、手術室プロトコル、及び専門家による助言に基づいて作成された。タスクは、患者関連行動、設備関連行動、コミュニケーション関連行動のいずれかに分類され、チームのパフォーマンスに応じて「はい」か「いいえ」のいずれかで評価された。OTAS は観察に基づく評価スケールであり、タスクは明確に実行された場合にのみ完了したとみなされる。

行動の評価は、Dickinson と McIntyre のチームワークモデルを応用して、協働モニタリング、コミュニケーション、協力、協調、協働リーダーシップといったチームワーク行動に基づいて行われる。行動パラメータは、Undre らによるインタビュー結果や他のチームワーク指標に基づいて設定され、Fletcher ら [8] が麻酔医の非技術的スキルの評価に用いたものも含まれる。チームパフォーマンスは、7 段階のリッカートスコアを用いてスコア化された (図 5) [34]。

良好なチームパフォーマンスの指標は最良の外科診療に関するガイドラインから導かれ、さまざまな行動パラメータと組み合わせて、手術室における特有のタスクの評価に用いられた。チームパフォーマンスに関するチェックリストは、一般外科と泌尿器科における日常的な手術室シナリオに重点を置いたものであり、術前・術後期間を含む手術のあらゆる側面に対応している。

RATING ANCHORS	BRIEF ANCHOR DEFINITION
6	Exemplary behaviour; very highly effective in enhancing team function
5	Behaviour enhances highly team function
4	Behaviour enhances moderately team function
3	Team function neither hindered nor enhanced by behaviour
2	Slight detriment to team function through lack of/inadequate behaviour
1	Team function compromised through lack of/inadequate behaviour
0	Problematic behaviour; team function severely hindered

Phase <sup>a</sup>	Stage 1	Stage 2	Stage 3
1. PRE	Pre-operative planning and preparation	Patient sent for to anesthesia	Patient set up to operative readiness
2. OP	Opening/ access to contact of target organ	Operative-specific procedure	From prepare to close to closure complete
3.POST	Anesthetic reversal to exit	Recovery and transfer	Feedback—self-assessment

### 図5 OTAS 評価スケール(上)と手術の場面・ステージ(下)

OTAS は5つの行動評価(コミュニケーション, 協調, 協力, リーダーシップ, 協働モニタリング: Communication・Coordination・Cooperation and Back up behaviour・Leadership・Team monitoring and Situation Awareness)と3つのサブチーム(外科医・麻酔科医・看護師)と3つの手術場面から  $5 \times 3 \times 3 = 45$  の評価ポイントがある[文献 34]。

### 評価システムの検証

OTAS のタスクチェックリストと行動評価スケールは, 専門家評価者と初心者評価者のそれぞれにおいて高い一貫性を示したことから, 高い構成概念妥当性を有していると考えられるが, この点についてはさらなる検討が必要である。

信頼性について Sevdalis らは, 専門家評価者と初心者評価者による, OTAS スケールを用いた手術室におけるチームワークの評価について検討を行った。その結果, 専門家評価者同士の間では高い相関が認められたが, 初心者評価者と専門家評価者との間の相関性は低かった。この結果から, OTAS スケールを用いた研修受講者の評価を行う前に, 初心者評価者に対して研修を行い, 彼らの学習曲線を求める必要があることが確認された。

研修での使用に関しては Undre らは, OTAS ツールを用いて, 結腸切除術, 前方切除術, 胃切除術, ハルトマン手術などの腫瘍切除術を含む泌尿器外科手術及び一般外科手術におけるチームワークパフォーマンスを評価した[35,36]。その結果, 手術室チームの全メンバーが, コミュニケーション行動において低いスコアを示した。また, 外科医の行動スコアは手術が終わりに近づくにつれて低下したが, この現象は, 外科医が手術の終盤に手術室を離れ, 後輩の外科医に閉創などの残りのタス

クの遂行を任せた場合に起こる可能性がある。また、機器のチェックや口頭での術式確認の怠りは頻繁に発生しており、手術の遅延や変更は全症例の約3分の2で発生していた。以上を総合すると、OTASは手術室における行動上の欠陥が起こる領域を特定する上で有用であるとともに、機器やタスクの欠陥が起こる領域を特定し、それらの特定と是正を可能にすると考えられる。

実用性については、OTASシステムの使用頻度はあまり高くないが、これはタスク特異的チェックリストが特定の手術にしか対応していないためである。さらに、本システムは活動中のリアルタイムな観察者による行動及びタスクの評価を必要とするため、実際の手術室環境で使用する評価方法としては、費用対効果と実用性の面で問題があると考えられる。OTASシステムは、手術室での評価のみに対応したシステムというよりも、システム全体の活動やチーム間の交流に関するより包括的な評価ツールとしての有用性が高いと考えられる。

## まとめ

手術室における行動の評価に特化した評価ツールとして、NOTSSとNOTECHSの2つの評価システムが代表的である。さらに、手術室におけるチームパフォーマンス行動の評価と手術タスクチェックリストを組み合わせた評価ツールであるOTASについてもここでは検討した。評価ツールの最終的な目的は、外科医が自身のノンテクニカルスキルに関するフィードバックと訓練を受けられるようにすることであり、その結果は、安全な手術の提供に直結する。NOTSSとNOTECHSの両スケールは妥当性が確立されたスケールであるとされているが、外科領域ではこれらのスケールの妥当性に関する実証研究については、NOTTSにおいては開始されつつあるが[23]、十分でとはいえず、今後はこの点についてさらなる検討を行う必要がある。OTAS尺度は、NOTSSやNOTECHSと比較して、評価対象とするチームワーク行動のカテゴリーが少ないが、周術期のタスクチェックリストが追加されているため、腫瘍症例などのより複雑な外科症例にも応用できる可能性がある。

いずれにしても評価ツールの開発が患者安全の目的ではなく、外科医個人や手術チームのノンテクニカルスキルを向上させ、より安全な手術の提供に繋がる評価ツールの開発と活用が求められる。

## 参考文献

- [1] Helmreich RL, Merritt AC, Wilhelm JA. The evolution of Crew Resource Management training in commercial aviation. *Int J Aviat Psychol* 1999;9:19-32.
- [2] Sharma B, Mishra A, Aggarwal R, Grantcharov TP. Non-technical skills assessment in surgery. *Surg Oncol*. 2011 Sep;20(3):169-77. doi: 10.1016/j.suronc.2010.10.001. Epub 2010 Dec 3.
- [3] Calland JF, Guerlain S, Adams RB, Tribble CG, Foley E, Chekan EG. A systems approach to surgical safety. *Surg Endosc* 2002;16:1005-14.
- [4] Healey AN, Sevdalis N, Vincent C. Measuring intraoperative interference from distraction and interruption observed in the operating theater. *Ergonomics* 2006;49:589-604.
- [5] Lingard L, Garwood S, Poenaru D. Tensions influencing operating room team function: does institutional context make a

difference. *Med Educ* 2004;38:

691 - 9.

[6] Lingard L, Reznick R, Espin S, Regehr G, DeVito I. Team communications in the operating room: talk patterns, sites of tension, and implications for novices. *Acad Med* 2002;77:232 - 7.

[7] Lingard L, Espin S, Whyte S, Regehr G, Baker GR, Reznick R, et al. Communication failures in the operating room: an observational classification of recurrent types and effects. *Qual Saf Health Care* 2004;13:330 - 4.

[8] Fletcher G, Flin R, McGeorge P, Glavin R, Maran N, Patey R. Anaesthetists' nontechnical skills (ANTS): evaluation of a behavioural marker system. *Br J Anaesth*

[9] Bogner M, editor. *Human error in medicine*. Hillsdale, NJ: LEA; 1994.

[10] Bogner M, editor. *Misadventures in Health care*. Mahwah, NJ: LEA; 2004.

[11] Gawande AA, Zinner MJ, Studdert DM, Brennan TA. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery* 2003;133:614 - 21.

[12] Gawande AA, Thomas EJ, Zinner MJ, Brennan TA. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. *Surgery*. 1999 Jul;126(1):66 - 75.

[13] Kable AK, Gibberd RW, Spigelman AD. Adverse events in surgical patients in Australia. *Int J Qual Health Care*. 2002 Aug;14(4):269 - 76.

[14] 医療事故情報収集等事業第 32 回報告書 [http://www.med-safe.jp/pdf/report\\_32.pdf](http://www.med-safe.jp/pdf/report_32.pdf)

[15] Flin R, Yule S, Paterson-Brown S, Maran N, Rowley D, Youngson G. Surgeons' non technical skills. *Surg News* 2005;4:83 - 5.

[16] Yule S, Flin R, Paterson-Brown S, Maran N, Rowley D. Development of a rating system for surgeons' non-technical skills. *Med Ed* 2006;40:1098 - 104.

[17] Sevdalis N, Davis RE, Koutantji M, Undre S, Darzi A, Vincent CA. Reliability of a revised NOTECHS scale for use in surgical teams. *Am J Surg* 2008;196:184 - 90.

[18] Undre S, Healey AN, Darzi A, Vincent CA. Observational assessment of surgical teamwork: a feasibility study. *World J Surg* 2006;30:1774 - 83

[19] Fletcher G, Flin R, McGeorge P, Glavin R, Maran N, Patey R. Rating non-technical skills: developing a behavioural marker system for use in anaesthesia. *Cogn Technol Work* 2004;6:165 - 71.

[20] O'Connor P, Flin R, Fletcher G. Methods used to evaluate the effectiveness of flight crew CRM methods in the UK aviation industry. *Hum Factors Aerospace Saf* 2002;2:235 - 55.

[21] Yule S, Flin R, Maran N, Rowley DR, Youngson GG, Paterson-Brown S. Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behaviour rating system. *World J Surg* 2008;32:548 - 56.

[22] Yule S, Rowley DR, Flin R, Maran N, Youngson G, Duncan J, Paterson-Brown S. Experience matters: comparing novice and expert ratings of non-technical skills using the NOTSS system. *ANZ J Surg* 2009;79(3):154 - 60

[23] Crossley J, Marriott J, Purdie H, Beard JD. Prospective observational study to evaluate **NOTSS** (Non-Technical Skills for Surgeons) for assessing trainees' non-technical performance in the operating theatre. *Br J Surg*. 2011 Jul;98(7):1010 - 20. doi: 10.1002/bjs.7478. Epub 2011 Apr 8

[24] Spanager L, Lyk-Jensen HT, Dieckmann P, Wettergren A, Rosenberg J, Ostergaard D. Customization of a tool to assess Danish surgeons' non-technical skills in the operating room.

Dan Med J. 2012 Nov;59(11):A4526

[25] 円谷 彰, 相馬 孝博, Yule Steven. [NOTSS-外科医に問われる手技以外のスキル]

医療安全およびチーム医療 外科チーム医療からみたノンテクニカルスキル. 臨床外科. [解説/特集]. 2013

2013.07;68(7):774-7.

[26] Klampfer R, Flin R, Helmreich RL, Hausler R, Sexton B, Fletcher G, et al. Enhancing performance in high risk environments: recommendations for the use of behavioral markers. Gottlieb Daimler and Karl Benz Foundation KollegGroup Interaction in High Risk Environments (GIHRE), <http://www.siaa.asn.au/simtect/2005/GIHRE2.pdf>; 2001 (assessed on 16.04 08).

[27] Flin R, Martin L, Goeters KM, Hörmann H-J, Amalberti R, Valot C, et al. Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills. Hum Factors Aerospace Safe 2003;3:97-119.

[28] O'Connor P, Hormann HJ, Flin R, Lodge M, Goeters KM. Developing a method for evaluating Crew Resource Management skills: a European perspective. Int J Aviat Psychol 2002;12:263-85.

[29] Moorthy K, Munz Y, Adams S, Pandey V, Darzi A. A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a simulated operating theatre. Ann Surg 2005;242:631-9.

[30] Moorthy K, Munz Y, Forrest D, Pandey V, Undre S, Vincent C, et al. Surgical crisis management skills training and assessment. Ann Surg 2006;244:139-47.

[31] Healey AN, Undre S, Vincent CA. Developing observational measures of performance in surgical teams. Qual Safe Health Care 2004;13(suppl.1):i33-40.

[32] Mishra A, Catchpole K, Dale T, McCulloch P. The influence of non-technical performance on technical outcome in laparoscopic cholecystectomy. Surg Endosc 2008;22:68-73.

[33] Mishra A, Catchpole K, McCulloch P. The Oxford NOTECHS System: reliability and validity of a tool for measuring teamwork behaviour in the operating theatre. Qual Saf Health Care 2009;18:104-8.

[34] Observational Teamwork Assessment for Surgery (OTAS) February 2011

[http://www1.imperial.ac.uk/resources/018F4A1D-5129-444E-96CF-04C524C2EA99/otas\\_manual.pdf](http://www1.imperial.ac.uk/resources/018F4A1D-5129-444E-96CF-04C524C2EA99/otas_manual.pdf)

[35] Sevdalis N, Lyons M, Healey A, Undre S, Darzi A, Vincent C. Observational teamwork assessment for surgery construct Validation with expert versus novice raters. Ann Surg; 249: 1047-051.

[36] Kim J, Neillipovitz D, Cardinal P, Chiu M, Clinch J. A pilot study using high-fidelity simulation to formally evaluate performance in the resuscitation of critically ill patients: the University of Ottawa Critical Care Medicine, high-fidelity simulation, and crisis resource management I study. Crit Care Med 2006;34:2167-4.

