

経験年数と患者情報の差異に着目した
理学療法士の臨床推論構造の
可視化に関する研究

総合情報学研究科
知識情報学専攻

インタラクションデザインの理論と実践

24M7111

高橋 可奈恵

論文要旨

近年、理学療法士（以下：PT）の臨床現場では、多くの患者情報の中で迅速な判断とリスク管理が求められており、その中でも臨床推論能力は重要な能力とされている。しかし、PTの臨床推論は暗黙知として扱われ、PT間で客観的に比較することが困難である。この課題に対して、本研究ではPTの臨床推論の思考過程を記述したテキスト（以下：アセスメント文）を構造化し、ネットワークで可視化することで客観的な比較が可能か検証する。加えて、経験年数や急性期から回復期の差異、提示する患者情報量と検査異常度の差異が臨床推論および共通解に及ぼす影響を検証する。これにより、臨床推論教育に資する知見を得るとともに、初学者PTに必要な支援の方向性を明らかにする。

本研究の提案手法では、アセスメント文を「問題点→対象部位→要因説明→根拠データ」の4項目へ構造化し、ネットワークとして生成する。ネットワークの可視化により、テキストのまま比較する場合に比べて、推論内容をより客観的に比較できることが示された。また定量的評価として、経験年数、大学教員によるパフォーマンス評価点、問題点数、ならびに正答に対する適合率・再現率・F値を算出し、相関分析を行った。その結果、経験年数の増加に伴い列挙される問題点数は減少し適合率は上昇する一方、再現率は低下傾向を示し、経験を重ねても見落としが生じ得る可能性が示唆された。

さらに、単一時点の比較にとどまらず、急性期・回復期の2時点の患者情報を提示し、初学者群と熟達者群の臨床推論構造にどのような差異が生じるかを検証した。具体的には、大腿骨転子部骨折術後7日目（以下：POD7）および術後30日目（以下：POD30）の同一模擬症例を提示し、両群のノード数・エッジ数およびネットワーク構造を比較した。その結果、初学者群ではPOD7からPOD30にかけてノード数が有意に減少し、集約ネットワークでは着目点が「筋力」に偏る参加者が多かった。一方、熟達者群ではノード数・エッジ数に有意な減少は認められなかったが、「筋力」に加えて「痛み」など複数の観点から問題点および検査に着目し、より網羅的に評価項目を列挙する傾向がみられた。

病期の差異に加え、提示する患者情報量と検査異常度の差異が臨床推論および共通解に及ぼす影響を検証した。専門家の合意による正解は仮定せず、参加者集団で列挙頻度が高い上位5位の「問題点と検査」の組み合わせを「共通解」と定義し、各参加者の適合率・再現率を算出した。模擬症例として、症例A（検査異常度：低い、情報制限）、症例B（検査異常度：低い）、症例C（検査異常度：高い）の3条件を設定し、初学者PT11名、熟達者PT11名を対象にアンケート調査を実施した。得られた回答から共通解を算出し、共通解に対する適合率・再現率について初学者群と熟達者群の群間比較をMann-WhitneyのU検定により行った。その結果、症例Cの心身機能における再現率のみ、熟達者群が有意に高かった。以上より、提示情報量が少なく、かつ検査異常度が低い条件では経験年数による差は大きく現れなかった一方、検査異常度が高い条件では、熟達者群の方が重要と共有されやすい「問題点」と「検査」の組み合わせを取りこぼしにくい可能性が示唆された。

本研究の結果から、初学者は重要な検査や問題点を見落としやすく、臨床推論が偏りやすいことが示唆された。そのため、臨床推論教育では、検査・問題点の取りこぼしや推論の偏りをリフレクションできる支援が必要である。今後は、見落としの低減と推論構造の可視化を目的としたシステムを構築し、臨床推論教育への貢献を目指す。

目次

1	序論	1
1.1	理学療法とは	1
1.2	臨床推論とは	1
1.3	理学療法士を取り巻く社会的背景	2
1.4	理学療法士の「質」について	3
1.5	理学療法学における臨床推論教育の課題	4
1.6	本研究で解決すべき課題	5
1.7	本論文の構成	6
2	関連研究	7
2.1	臨床推論教育の動向	7
2.2	臨床推論の知識共有に関する研究	7
2.3	知識工学分野における推論プロセスの可視化に関する研究	8
2.4	意思決定プロセスと認知バイアスに関する研究	10
2.5	本研究の立ち位置	11
3	実験デザイン	12
3.1	本研究のリサーチクエスション	12
3.2	リサーチクエスションに対する実験デザイン	13
4	アセスメント文の構造化に関する検証	17
4.1	アセスメント文を構成する項目の設定	17
4.2	アセスメント文の構造を表すネットワーク表現	17
4.3	検証	18
4.4	検証結果	19
4.5	考察	23
4.6	本章のリミテーション	23
4.7	本章の結論	24
5	病期の差異による初学者群・熟達者群の臨床推論構造の比較検証	25
5.1	ネットワークの項目の設定	25
5.2	実験目的	27
5.3	実験方法	27
5.4	検証と結果	28
5.4.1	検証1：記述統計と代表例のネットワークの比較	28
5.4.2	検証2：初学者群・熟達者群の群間比較	32
5.4.3	検証3：術後経過による変化（POD7とPOD30の比較）	32
5.4.4	検証4：集約ネットワークによる比較	35
5.5	考察	38

5.6	本章のリミテーション	38
6	提示情報量と検査異常度の差異が臨床推論ネットワークおよび共通解に及ぼす影響の検証	40
6.1	「共通解」に関する研究と定義	40
6.2	ネットワークの項目の設定	40
6.3	実験目的	41
6.4	実験方法	41
6.5	検証と結果	43
6.5.1	検証1：ICF（心身機能・活動・参加）における問題点列挙の比較	43
6.5.2	検証2：共通解との適合率・再現率	49
6.5.3	検証3：集約ネットワークの比較	53
6.6	本章の考察	57
6.7	本研究の展望	58
7	結論	60
8	謝辞	62

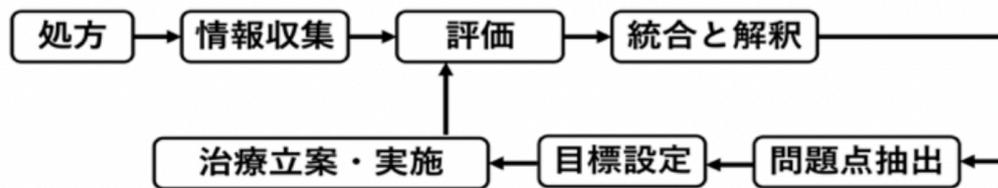


図 1.1: 理学療法プロセスの枠組み

1.3 理学療法士を取り巻く社会的背景

本邦では少子高齢化の進行に伴い、理学療法士を取り巻く社会的背景はこの30年で大きく変化し、臨床推論能力の重要性は高まった。2000年度の診療報酬改定において、リハビリテーション医療の拡充を目的とした「回復期リハビリテーション病棟入院料」が新設された。また、同年には介護保険制度が始まり、高齢化社会への対応が本格化した時期であった。2000年以前では、病院での長期入院が前提とされ、在院日数は長く、急性期治療後も回復期および慢性期が同一病棟内に混在する形で医療が提供されていた。また当時のリハビリテーション医療は経営的観点からも厳しく、公的病院や温泉地などの郊外における運営が中心であった。その結果、適切なリハビリテーション治療を十分に受けることなく、生活期へと移行する患者も少なくなかった [19]。

こうした背景の下、2000年度の診療報酬改定において「回復期リハビリテーション病棟入院料」が新設された。本改定では、医療の質向上と効率化、医療機関の機能分担と連携の促進、入院医療の機能分化といった基本方針が明示されている³。その中で「回復期リハビリテーションの充実」が重要な施策として位置づけられ、「寝たきり状態の患者の発生を防止するための回復期リハビリテーションを行う病棟入院料の新設」が示された。すなわち、制度創設当初から、回復期リハビリテーションには「必要な時期に集中的な治療を行い、機能回復を促進し、寝たきり化を防止する」という明確な目的が付与されていた。

さらに2006年の診療報酬改定では、理学療法・作業療法・言語聴覚療法が再編され、「疾患別リハビリテーション料（脳血管疾患等・運動器・呼吸器・心大血管）」が新設された。これにより、理学療法の算定は従来の「提供した手技」単位の評価から、対象疾患群および体系に基づく評価へと整理された⁴。これは、これまでの1単位20分での一定の料金体系ではなく、疾患ごとに細かく価格設定がなされた。例えば、脳血管疾患であれば1単位2450円に対し、運動器疾患は1850円であり、脳血管疾患の重要性がクローズアップされた結果になった。さらに、本改定では各疾患別リハビリテーション料に算定日数の上限が設定され、医学的有用性および医療資源の効率的活用の観点から、リハビリテーションを漫然と継続するのではなく、限られた期間内で目標を設定し、治療効果を判断することが制度的に求められるようになった。

その後、2018年の診療報酬改定では、集中治療室等における「早期離床・リハビリテーション」を促進する加算が新設された。本加算では、多職種チームにより計画を作成し、入

³国立社会保障・人口問題研究所 <https://www.ipss.go.jp/publication/j/shiryuu/no.13/data/shiryuu/iryuu/307.pdf>(2026/2/1 確認)

⁴厚生労働省 <https://www.mhlw.go.jp/shingi/2006/02/dl/s0215-3u.pdf>(2026/2/1 確認)

2-2. 年齢区分

	n=862		n=968		n=700		n=342		n=301		n=132	
	病院に勤務						関連する介護保険施設等に勤務					
	理学療法士		作業療法士		言語聴覚士		理学療法士		作業療法士		言語聴覚士	
	人数	割合	人数	割合	人数	割合	人数	割合	人数	割合	人数	割合
21～30歳	7,846	56.3%	4,350	53.1%	1,470	47.3%	1,106	47.1%	570	43.1%	133	42.2%
31～40歳	4,022	28.9%	2,784	34.0%	1,181	38.0%	791	33.7%	522	39.5%	121	38.4%
41～50歳	1,503	10.8%	847	10.3%	322	10.4%	342	14.6%	197	14.9%	51	16.2%
51～60歳	509	3.7%	186	2.3%	110	3.5%	99	4.2%	30	2.3%	8	2.5%
61～70歳	46	0.3%	22	0.3%	22	0.7%	10	0.4%	4	0.3%	2	0.6%
合計	13,926	100.0%	8,189	100.0%	3,105	100.0%	2,348	100.0%	1,323	100.0%	315	100.0%

図 1.2: 平成 28 年時の年齢分布 [18]

院後早期から治療を行った場合に、14日を限度として評価できる枠組みが明記されている。2022年改定では、本加算の対象が救命救急入院料や高度治療室（HCU）等へと拡大された⁵。

以上のように、診療報酬制度の変遷を通じて、理学療法士を取り巻く医療環境は大きく変化してきた。2000年代以降、日本の医療制度は、診療報酬改定や地域包括ケアの推進、在院日数短縮と機能分化、多職種連携およびアウトカム重視といった方針を軸に、臨床現場の意思決定をより迅速かつ説明可能なものへと転換してきた。その結果、理学療法士には、手技や運動処方熟練にとどまらず、限られた時間と情報の中で問題を同定し、優先順位を付け、治療の根拠と見通しを言語化する臨床推論能力が強く求められるようになってきている。

1.4 理学療法士の「質」について

近年、「理学療法士の質の低下」が懸念されている。現在、日本の理学療法士の数は増加傾向にあり、国際的な比較でもその伸び率は大きい。著者の調べによると、10万人あたりの理学療法士の数は、2013年時点で豪州113人、日本79人、米国64人だったが、2019年には豪州141人、日本167人、米国88人と、日本は倍増している。これに伴い、日本の理学療法士の平均年齢は33歳となり、豪州38歳、米国41歳と比較しても、理学療法士の若年化が進んでいる。この現象の背景として、理学療法士学校養成施設の規制緩和や規制改革が指摘されている。これらの規制緩和は、リハビリテーションに対する社会的要請や回復期リハビリテーション入院料の創設、さらには介護保険制度の施行によるものであり、2000年頃からの高齢化問題に対応するため、専門職を急速に増やす必要があったと考えられる [18]。

質の低下の要因として、若い理学療法士の増加は歓迎される一方で、その急増が理学療法士の質の低下を招いている可能性があることは無視できない。新卒の理学療法士（以下、「初学者理学療法士」）の数が急増する中、指導者が不足しており、ピラミッド型の人員配置が現場の教育体制に大きな負担を与えていると考える。

特に、高齢化が進む現代は単に高齢者が増えるという理解ではなく、平均寿命の伸びによって高齢者の年齢が高くなっていることにも目を向けなければならない。現代の医療は年齢が進むことにより、原疾患が完治する前に他の病気を罹患する複合疾患という問題に

⁵厚生労働省 <https://www.mhlw.go.jp/file/06-Seisakujouhou-12400000-Hokenkyoku/0000198532.pdf> (2026/2/1 確認)

直面させられている。一人の患者に複数の疾患が併存しているため、結果的に高齢患者の増加は理学療法士に対して、運動器、神経疾患、循環器、呼吸器など多岐にわたる知識が求められていることに繋がっている。このため、初学者理学療法士に対する卒前後教育の負担はさらに増大している。規制緩和によって入学定員が急増し、国家試験合格率が低下する中で、卒業直後の初学者理学療法士のレベルは依然として「多くの助言を要する」状態であり、養成の質の低下が指摘されている [18]。また、初学者理学療法士の急増によって、指導者と初学者の比率が不均衡になっていることも大きな課題である。現状の理学療法の教育において、養成校で技能を完成させることは難しく、同時に教育者ではない現場スタッフが業務を行いながら教育効果をだすことは困難であると言わざるを得ない。

実際に、卒後教育に関する調査結果では、施設で行われている教育は指導者の裁量に依存する部分が大きく、卒後教育の質は指導者によって大きく異なる可能性が示唆されている [16]。さらに、少人数の施設では、充実した卒後教育システムが整っていない可能性も指摘されている [16]。卒後の初学者理学療法士の指導経験がある理学療法士に対して、指導している理学療法士がどのような能力を獲得したときに自立したと感じるかのインタビューが行われた研究では、「理学療法実施上の必要な知識」「理学療法士としての技術」「コミュニケーション技術」「態度」「自己教育力」「自己管理能力」「臨床思考能力」の中では、「臨床思考能力」が必要だと回答する理学療法士が最も多かった [34]。したがって、卒後の理学療法士が自立した理学療法士への成長を促すためには、臨床現場での臨床推論教育が必要とされる。

1.5 理学療法学における臨床推論教育の課題

近年の診療報酬改定や在院日数短縮、機能分化、多職種連携およびアウトカム重視といった制度変化により、理学療法士に求められる専門性は大きく変容してきた。このような状況下で臨床推論教育の重要性が増している理由は、単に学生や若手理学療法士へ高い知識を獲得されるためでなく、専門職としての判断の質を担保する仕組みが、現場および制度側から強く要請されるようになったためである。制度がアウトカムや説明責任を重視する方向へ進むにつれ、理学療法は「何を実施したか」ではなく、「なぜその治療を選択したのか」、「それによってどのような変化が見込まれるのか」といった判断過程を問われるようになっていく。

しかし、臨床推論教育が十分に行われない場合、臨床実践は既存パターンの模倣に偏りやすく、診療記録は定型文の羅列となり、評価行為そのものが目的化する危険性がある。このような状況では、症例ごとの違いに応じて治療を判断する力が育ちにくく、結果として「誰が実施しても同じ」「理学療法でなくても代替可能である」といった評価に近づくことが懸念される。これは個々の理学療法士の努力不足ではなく、業務の過密化の中で暗黙知の継承に依存し続けてきた結果であると考えられる。臨床推論教育は、この暗黙知を構造化・可視化することで、判断過程の再現性と説明可能性を臨床現場にもたらす役割を担う。

さらに、安全性の観点からも臨床推論教育の重要性は高い。高齢化の進行や併存疾患の増加により、理学療法の対象は年々複雑化しており、リスク管理、禁忌の判断、急変兆候の早期発見、疼痛や心理社会的要因への対応など、誤った判断が患者に重大な不利益をも

たらず場面が増加している。ここで理学療法士に求められるのは治療の正解を知っていることではなく、限られた情報から仮説を立て、危険な可能性を除外しつつ、説明と合意の下で治療の強度や優先順位を調整する能力である。臨床推論教育は、こうした判断を個人の経験や勘に依存させるのではなく、他職種で共有可能な形式に構造化するための訓練でもある。

しかしながら、現行の理学療法養成校教育は、必ずしもこの要請に十分応えられているとは言い難い。理学療法養成校での教育は、厚生労働省が定める「理学療法作業療法学校養成施設規則」に基づいて実施されており、カリキュラムは基礎分野、専門基礎分野、専門分野に区分されている。専門分野には「理学療法評価学」が設けられているものの、その教育目標は理学療法評価に関する知識と技術の習得に主眼が置かれており、臨床推論に関する具体的な教育内容は明確に規定されていない[23]。一部の養成校では「臨床推論」や「理学療法診断論」といった科目を設ける例も見られるが、養成校の教育において臨床推論を体系的に学習できる環境が十分に整備されているとは言えない。

また、初学者理学療法士を対象とした調査では、姿勢や動作の観察・分析、および得られた情報の統合と解釈に関する自己評価が低いことが報告されており[35]、多くの初学者が臨床推論を要するプロセスに困難さや苦手意識を抱いている実態が示されている。このような養成校教育と臨床現場の間に存在するギャップを踏まえると、臨床推論教育は、知識を教えることにとどまらず、知識を用いて判断を行う回路を形成するための中核的教育であると言える。臨床推論教育を通じて、教育と実践の断絶を縮小し、理学療法の質と安全性を担保する仕組みを構築することが、今後ますます重要になると考えられる。

1.6 本研究で解決すべき課題

本研究で解決すべき主たる課題は、理学療法士の臨床推論が暗黙知化しており、再現性および説明可能性が十分に担保されておらず、客観的に臨床推論を比較しづらい点にある。暗黙知は哲学者のマイケル・ポランニーによって提唱されたもので、形式知と対称関係を持つ[29]。勘やコツのような領域として臨床推論が理解されており、観察と経験により培われるものであるという考え方が、現場の一定理解となっている。そのためか、1.5章で述べたように、理学療法教育の観点から見ると、臨床推論は体系的かつ構造的に教育されているとは言い難い。理学療法養成校では国家試験対策が重視される傾向にあり、カリキュラム上、「理学療法評価学」は設けられているものの、臨床推論を体系的に学習する教育設計は標準化されていない。一方で、1.4章で述べたように、臨床現場では臨床推論能力が強く求められているにもかかわらず、養成校教育と臨床現場の間には依然としてギャップが存在している。さらに、臨床現場では初学者理学療法士の急増により、初学者と指導者の比率が不均衡となり、十分な臨床教育体制が整っていない可能性がある。本研究では、臨床現場で求められる臨床推論を構造化することにより、養成校および臨床現場の双方において、再現性の高い標準的な臨床推論教育を担保しやすくすることを目指す。

臨床推論が構造化されていない場合、理学療法士の思考は個人的な経験や勘に依存しやすく、暗黙的な判断に陥りやすい。その結果、理学療法士個人間での比較や知識共有が困難となるという課題が生じる。初学者理学療法士と熟達者理学療法士の臨床推論の違いを明

らかにすることで、それぞれが陥りやすい推論構造を把握できる可能性がある。また、患者の病期や患者の情報量、検査の結果によってによって臨床推論がどのように変化するのかを、客観的に比較できると考えられる。特に、手術直後の急性期と回復期の差異の影響、ならびに模擬患者に提示する情報量や症例難易度を段階化した場合の影響を比較することにより、推論構造の特徴や偏りを明らかにし、初学者理学療法士に対する臨床推論教育への知見を得ることを本研究の目的とする。

1.7 本論文の構成

本論文の構成は、以下の章からなる。1章では、理学療法士の現状の背景と臨床推論教育の問題点について整理した。2章では、理学療法の臨床推論教育の関連研究と思考を構造化する関連研究を紹介する。3章では、リサーチクエスチョンを提起し、それに対応する実験デザインを述べる。4章では、統合と解釈文の構造化に関する検討について述べる。5章では、臨床推論を構造化したネットワークを用いて、術後7日目から30日目にかけての病期の差異に伴う、初学者群と熟達者群のネットワークの比較について述べる。6章では、実験参加者間での「共通解」を定め、初学者群と熟達者群の適合率や再現率、またネットワークを比較し、それらの結果から展望を述べる。7章では本研究の結論を述べる。

2 関連研究

2.1 臨床推論教育の動向

近年の理学療法士における臨床推論教育では、臨床現場での担当患者を題材にした症例レポートや症例検討会が主に各施設で実施されている。日本理学療法士協会の生涯学習制度における前期研修では、eラーニングを通じた座学が提供されており、その中の「理学療法の基礎」という講座には「クリニカルリーズニング」に関する講義が設けられている¹。また、医学教育を中心に導入されている臨床技能評価法である客観的臨床能力試験 (Objective Structured Clinical Examination : 以下 OSCE) が理学療法教育にも応用され、理学療法版 OSCE が 2007 年度に開発されている [31]。2016 年には理学療法士学生を対象に、映像とグループリフレクションを取り入れた「考える OSCE-R」がより高度な臨床推論を行うための方法として開発され、真正性の高いパフォーマンス評価を示したことが報告されている [27]。しかし、症例レポートなどのテキストを用いた臨床推論の評価手法は未だ確立されていない [17]。

他にも理学療法のみならず、医学教育においての不確実な状況下での臨床推論を評価するために用いられる、スクリプトコンコーダンステスト (Script Concordance Test : 以下 SCT) がある [2]。SCT では、診断・検査・治療などを仮説を設定した短い文章で示された模擬症例に対して、模擬症例の新たな情報を提示し、その仮説の確からしさを段階尺度で回答させる形式の評価方法である [9] (図 2.1)。SCT の特徴は単一の正解を仮定せず、専門家の最頻回答を満点とし、他の選択肢も選択した割合に応じて得点化されるため、臨床における判断を一定に反映した評価が可能となる。しかし、SCT は「模擬患者の新情報の追加に伴う仮説の更新」という判断を測定対象とするため、医学部学生がどこの情報に着目し、どのような推論過程を経て、その判断にいったのかが表出されない。また医学部生と専門家の推論過程の差異が汲み取りにくく、特徴や見落とししやすい観点を抽出する方法は未だ確立されていない。

2.2 臨床推論の知識共有に関する研究

臨床推論能力の向上を目的とした研究としては、知識共有の構築に関する研究が行われている [33], [32], [30]。吉田らは、理学療法プロセスの「②情報収集」に着目し、情報収集段階で着眼点を提示することで、情報の見落としを防ぎ、情報の分析と治療すべき問題点の抽出が促進されるかを検証した [33]。この研究では、実験協力者に模擬患者症例情報の確認から治療すべき問題点の抽出を行わせる際、着眼点の提示の有無を操作することで問題点の列挙と情報分析の記述が向上するかを検証している。その結果、情報収集時の着眼点を提示した群では、治療すべき問題点を列挙することができていること、ならびに情報分析時の記述において余分な情報が減少したことが確認された。しかし、提案手法は論理的な解釈に至る前段階の支援にとどまっており、論理性を高めるには文章同士のつながりや主張している仮説をより詳しく分析する必要がある。

¹公益社団法人日本理学療法士協会:生涯学習制度 前期研修カリキュラム https://www.japanpt.or.jp/pt/lifelonglearning/asset/pdf/zenki_curriculum_20221109.pdf(2025/1/20 確認)

(a) Judgment type: Diagnosis

A 58-year-old woman presents to the emergency department with a two-week history of intermittent vertigo. She feels well between episodes.

If you were thinking of:	And then you find:	This diagnosis becomes:				
		-2	-1	0	+1	+2
Q1. Benign paroxysmal positional vertigo	Episodes of vertigo last 30 minutes	-2	-1	0	+1	+2
Q2. Transient ischemic attacks	History of hypertension	-2	-1	0	+1	+2
Q3. Meniere's syndrome	Recent surgical removal of a skin lesion	-2	-1	0	+1	+2

図 2.1: SCT の模擬症例の例 [9]

宮本らは、動作分析テキストに現れる実践知の可視化を行った [30]。理学療法士は、対象者の日常生活動作の改善を目的に理学療法評価の一つの方法として動作分析を行う。動作分析とは、取得したデータや目視による動作観察結果をもとに、動作における問題点を分析する過程である。この分析は視覚的な観察に基づくため、主観性が高く、再現性が低いという課題がある。そのような課題があるため、動作分析の文章には実践知が多く含まれていることが特徴である。こうした実践知を共有するために、宮本らは動作分析テキストを理学療法に関する知識をもとに分類し、因果関係のある要素を紐付けてネットワーク形式で可視化を行った。その結果、ネットワークのバランスや理学療法知識の因果関係による紐付けの観点から、動作分析における実践知の有無が論理構成に与える影響を把握できる可能性が示唆された。しかし、動作分析は理学療法のプロセスにおいて理学療法評価の一部を担うものであり、理学療法プロセスの「②情報収集」から「④統合と解釈」に位置づけられるものである。そのため、臨床推論能力が特に求められる場面で、初学者の理学療法士が苦手とする「④統合と解釈」に対する支援が十分ではない。

2.3 知識工学分野における推論プロセスの可視化に関する研究

臨床推論を構造化するために、医療分野以外での知識工学分野における推論プロセスの可視化に関する研究を紹介する。Abraham らは、知識を「概念」と「概念同士のつながり」として捉え、概念をノード、つながりをエッジとして表現することで、学習者の知識構造を可視化・評価する枠組みを示している [1]。多肢選択式問題の解答データをもとに、学習者がどの概念を理解しているかだけでなく、概念同士をどのように関連づけているかを推定し、ネットワークを生成する。そして、教員が想定するネットワークと学習者のネットワークを比較し、「どれだけ似ているか」を数値化することで理解の程度を評価する [1]。また、参照マップ側で重要な概念を見つけるために中心性などのネットワーク指標を利用し、

示され、推論マップを通じてそれらのバイアスを視覚的に把握し、創造的思考過程の中で意識的に修正・排除できる可能性が示唆されている。

2.4 意思決定プロセスと認知バイアスに関する研究

1.1章で述べたように、臨床推論は、臨床における思考および意思決定のプロセスである。また1.3章で述べたように、2000年以降、診療報酬改定の推進や、在院日数短縮と機能分化、多職種連携、アウトカム重視といった方針を背景として、臨床現場の意思決定は、より迅速で説明可能なものへと転換してきた。一方、業務の過密化が進む中で、臨床推論は既存パターンの模倣に偏りやすく、診療記録は定型文の羅列となりやすい。その結果、本来は意思決定を支えるための評価行為が、目的そのものになってしまう危険性も指摘されている。

意思決定プロセスに関する研究は、心理学および行動経済学を中心に発展してきたが、近年では医療分野においても、診断エラーや臨床判断を理解するための枠組みとして重要性が高まっている。特に、瞬間的に自動的に生じる直感の心理（システム1）と、分析的な熟慮の心理（システム2）からなる二重過程理論（Dual Process Theory）は、人間の判断を説明する代表的な理論である [6][12]。Evansらは、自動的に迅速な処理がまずデフォルトの反応を生成し、必要に応じて、高次の推論過程がこれに介入すると仮定する。両者の本質的な違いは、システム2処理が仮説的思考を可能にし、ワーキングメモリに強く依存すると整理している [6]。

医療意思決定研究では、二重過程理論を基盤として診断推論や診断エラーの理解が進められてきた。Croskerryは、医療者の診断過程が直感の心理と熟慮の心理の往復によって成り立っていることを示し、診断エラーを単なる知識不足ではなく認知プロセスの問題として捉える枠組みを提示した [4]。診断エラーは複数の要因によって生じるが、その主要因として認知バイアスを含む思考過程の問題が挙げられている。また、認知過負荷や疲労、時間的制約などの要因が分析的思考を阻害し、診断エラーを助長することが指摘されている [4]。

このような認知プロセスに基づく診断エラーの理解は、海外のみならず国内の医学分野においても展開されている。本邦における医療訴訟判例の分析から、「診断の見逃し」や「診断の遅れ」が訴訟原因の上位を占め、その多くが医師個人の能力不足ではなく、認知バイアスや想起エラーといった認知エラーに起因することを示している [22]。また、診断エラーの多くは、正しい診断仮説が想起されない場合や、認知バイアスによって妥当な仮説が推論過程の途中で却下される場合に生じると整理されている。これらのバイアスは、迅速かつ自動的に作動するシステム1で生じやすく、推論過程の早期に正確な診断の可能性を棄却してしまう早期閉鎖（premature closure）を引き起こす。この知見は、二重過程理論に基づく診断エラー理解と整合的であり、診断エラーを意思決定プロセス全体の問題として捉える視点の重要性を示している [22]。

Croskerryらは、認知バイアスの低減は一度きりの教育介入によって達成されるものではなく、「省察」「ルール化」「認知的強制戦略」といった仕組みを通じて、日常の意思決定の中で徐々に形成されると述べている [5]。認知バイアスに対する対策として、Croskerryらは認知的強制戦略を含むデバイアス手法を提案しており、教育的介入とシステムの支援の

両面から診断エラー低減を図る必要性を指摘している [5].

2.5 本研究の立ち位置

これまでの関連研究から、臨床推論教育においては、症例レポートや OSCE, SCT などを通じて判断結果を評価する試みや、着眼点の提示や知識共有によって情報収集や分析を支援する試みが行われてきたことが示されている。また、知識工学分野においては、知識構造や推論過程をネットワークやマップとして可視化し、学習者の理解状況や思考の特徴を把握・評価する枠組みが提案されてきた。

一方で、理学療法における理学療法プロセスの「④統合と解釈」では、評価結果をどのように関連づけ、どのような思考過程を経て判断に至ったのかといった推論過程そのものが十分に表出・共有されていないという課題が残されている。既存の評価手法や支援手法では、判断の妥当性や結果を捉えることはできるものの、初学者と熟達者の推論構造の比較や、見落としやすい観点、推論パターンの偏りを把握することは容易ではない。

さらに、意思決定プロセスと認知バイアスに関する研究からは、臨床推論が直感的判断と分析的判断の相互作用によって成り立っており、経験の蓄積が必ずしも安全な判断を保証しないことが示されている。このことは、臨床推論の支援において、判断手順を標準化するだけでは不十分であり、どの情報をどのように結びつけて解釈しているのかという思考過程を可視化し、省察や再検討を促す枠組みが必要であることを示唆している。以上を踏まえ、本研究では、臨床推論構造を比較可能な形式へ変換し、臨床推論教育に対する知見を得るとともに、初学者理学療法士に必要な支援を明らかにすることを目的とする。

3 実験デザイン

3.1 本研究のリサーチクエスチョン

1章で述べたように、近年の少子高齢化を背景として、理学療法士の臨床現場では大量かつ多様な患者情報を踏まえた迅速な判断とリスク管理が求められており、その中でも臨床推論能力は重要な能力とされている [34]。しかし、理学療法士の臨床推論は暗黙知として扱われやすく、思考過程の再現性および説明可能性が十分に担保されていないため、理学療法士間で客観的に比較することが困難である。その結果、経験年数に伴う臨床推論の違いが明確になりにくく、「熟達」の特徴を形式的に表現することは容易ではない。さらに、患者情報量や検査所見の異常の程度が臨床推論に与える影響についても、検証の枠組みは十分に整備されていない。これらの課題は、養成校および臨床現場における臨床推論教育の体系化を妨げ、初学者が臨床推論に苦手意識を抱く一因となっている。

以上を踏まえ、本研究では、臨床推論の思考過程を比較可能な形式へ変換し、臨床推論教育に対する知見を得るとともに、初学者理学療法士に必要な支援を明らかにすることを目的とする。この目的を踏まえて、「臨床推論構造をネットワークで可視化することにより、経験年数および患者情報の差異に応じて推論構造はどのように異なるのか」をリサーチクエスチョン（以下：RQ）として提起する。また、RQを解決するため、以下の3つのサブリサーチクエスチョン（以下：SRQ）へ細分化し、段階的に検証を進める。

RQ：臨床推論構造をネットワークで可視化することにより、経験年数および患者情報の差異に応じて推論構造はどのように異なるのか

- SRQ1「アセスメント文を構造化し、ネットワークで可視化することによって、理学療法士間の臨床推論を客観的に比較することは可能か」
- SRQ2「急性期・回復期の2地点の患者情報提示により、初学者・熟達者の臨床推論ネットワークはどのような差異が見られるか」
- SRQ3「提示する患者情報量と検査異常度の差異により、初学者・熟達者の臨床推論ネットワーク、ならびに共通解はどのように異なるのか」

RQと細部化したSRQの関連図は図3.1に示す。SRQ1を解決するため、臨床推論能力が最も求められる理学療法プロセスの「⑤統合と解釈」における記述文を「アセスメント文」と定義し、このアセスメント文を構造化を行い、ネットワークで可視化する手法を提案する。従来の文章表現と比較して、臨床推論をより客観的に比較可能な形式へ転換できるかを検証する。具体的には、治療対象となる問題点の列挙数や、正答とした問題点との適合率・再現率などの指標を用い、経験年数による比較を行うことで、初学者および熟達者に特徴的な傾向を明らかにする。

SRQ2では、RQ1で構築した臨床推論ネットワークの枠組みを用い、単一時点での群間比較にとどまらず、急性期および回復期の2時点における推論構造の差異を検証する。病期に伴う推論構造の差異を分析するとともに、初学者・熟達者それぞれの特徴や着目点の違いを明らかにし、臨床推論教育への具体的示唆を提案する。

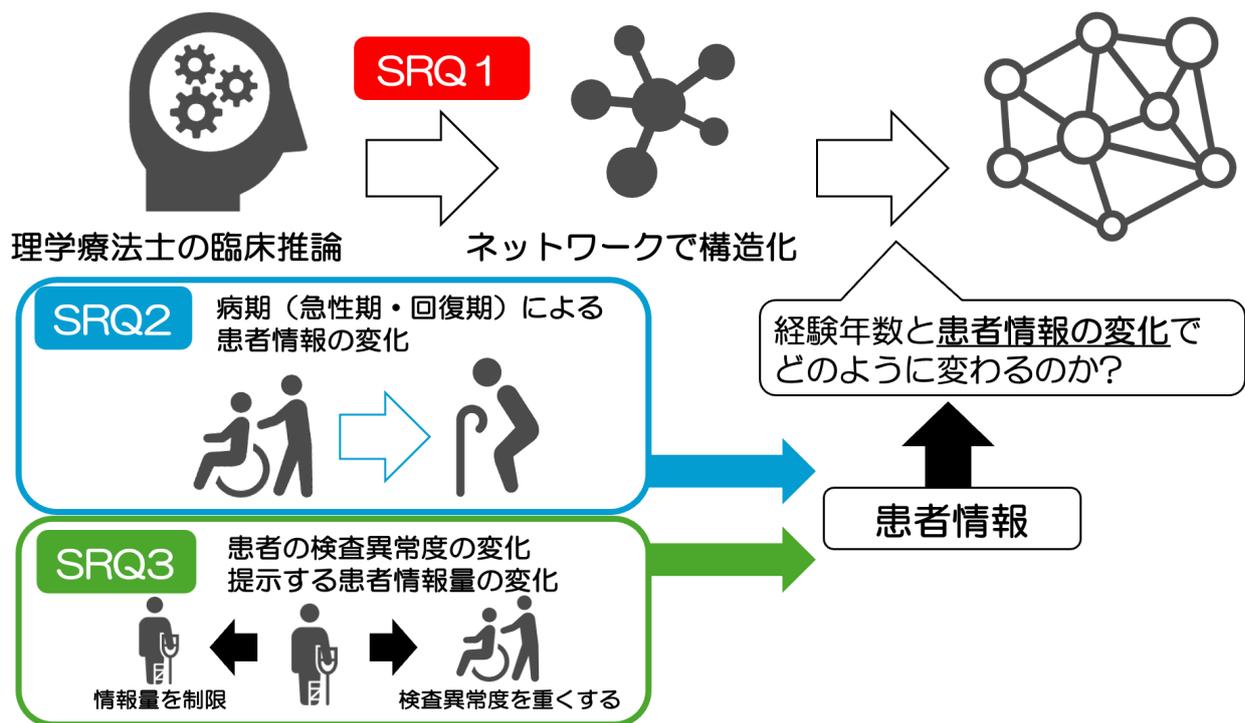


図 3.1: RQ の関連図

SRQ3 では、病期の差異だけでなく、模擬患者に提示する情報量と検査異常度を段階付けることで、理学療法士の推論構造および共通解がどのように異なるのかを検証する。提示する患者情報量を制限する症例や検査異常度を軽度から中等度なレベルへ段階的に設定した症例を用い、初学者群と熟達者群の推論を比較する。さらに、実験参加者間で共通して選択された問題点と検査の組み合わせを「共通解」として算出し、2.4 章で述べた認知バイアスの観点から、推論の収束過程および見落としが生じる特徴を分析する。

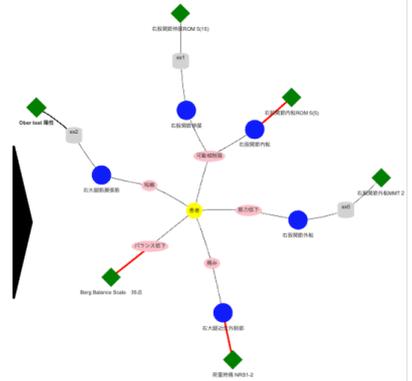
3.2 リサーチクエスチョンに対する実験デザイン

- 第 1 実験：アセスメント文の構造化に関する検証

第 1 段階目の実験 (以下：第 1 実験) では、SRQ1 の「アセスメント文を構造化し、ネットワークで可視化することによって、理学療法士間の推論を客観的に比較することは可能か」を明らかにする。この問いに対して、アセスメント文を構造化し、ネットワークとして可視化する手法を提案する (図 3.2)。具体的には、アセスメント文を「問題点→対象部位→要因説明→根拠データ」の 4 項目へ構造化し、ネットワークとして生成する。また、経験年数 1~5 年目の PT が作成したアセスメント文 20 件を対象とし、経験年数、大学教員によるパフォーマンス評価点、問題点の数、ならびに問題点の正答に対する適合率・再現率・F 値を算出する。定量的評価として経験年数を基準に、問題点の数、問題点の正答に対する適合率・再現率・F 値で相関分析を行う。また定性的評価にて、ネットワークの可視化が有効か検討する。



第1層 「問題点」	第2層 「対象部位」	第3層 「要因説明」	第4層 「根拠データ」
筋力低下	右股関節外転	右荷重歩容期、立脚中期においてデュシェンヌ様の跛行	右股関節外転MMT2
可動域制限	右股関節内転		右股関節内転ROM 5(5)
可動域制限	右股関節伸展	右kneeで大腿直筋の短縮があるため、股関節伸展制限が認められる	右股関節伸展ROM 5(15)
痙攣	右大腿筋線張筋	大転子遠位外側を皮切し、大腿筋線張筋を切離しているため	Ober test 陽性
痛み	右大腿近位外側部		疼痛特異 NFS 1-2
バランス低下			Berg Balance Scale 35点



①アセスメント文を自由記載で収集

②各項目に分けて構造化

③臨床推論ネットワークを生成

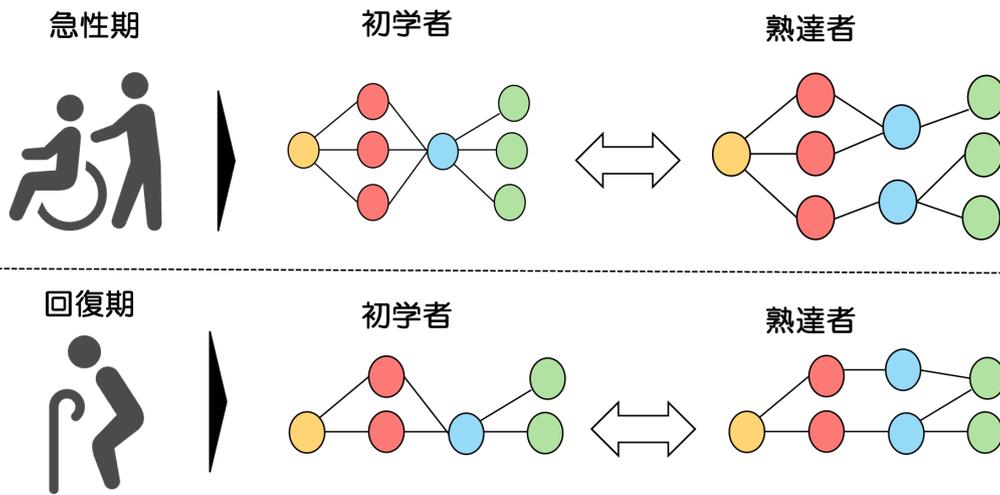
図 3.2: 第1 実験：アセスメント文の構造化に関する検証

- 第2 実験：病期の差異による初学者群・熟達者群の臨床推論構造の比較検証

第2段階目の実験（以下：第2 実験）では、第1 実験で臨床推論を構造化したネットワークを用いて、SRQ2の「急性期・回復期の2地点の患者情報提示により、初学者・熟達者の臨床推論ネットワークはどのような差異が見られるか」という問いを明らかにする。第2 実験では、第1 実験で用いた記述文は用いず、「問題点（心身機能）→身体部位→検査」の3つを質問項目とした選択式アンケート調査を作成する。実験対象者は初学者理学療法士10名・熟達者理学療法士10名に設定する。また、単一時点の比較にとどまらず、急性期・回復期の2時点の患者情報を提示し、初学者群と熟達者群の臨床推論ネットワークにどのような差異が生じるかを検証する（図3.3）。具体的には、大腿骨転子部骨折術後7日目（以下：POD7）および術後30日目（以下：POD30）の同一模擬症例を提示し、両群のノード数およびエッジ数を算出して比較する。POD7とPOD30で初学者・熟達者群でノード数およびエッジ数に差はあるのか、Mann-WhitneyのU検定を行う。加えて、POD7からPOD30にかけてノード数・エッジ数が有意に減少するかを検証するため、参加者全体、初学者群、熟達者群の3条件について、対応のあるWilcoxon符号付順位検定を行う。また初学者群・熟達者群それぞれの集約ネットワークを生成し、各群での特徴や偏りを明らかにする。

- 第3 実験：提示情報量と検査異常度の差異が臨床推論ネットワークおよび共通解に及ぼす影響の検証

第3段階目の実験（以下、第3 実験）では、SRQ3である「提示する患者情報量と検査異常度の差異により、初学者・熟達者の臨床推論ネットワーク、ならびに共通解はどのように異なるのか」という問いを明らかにする。第2 実験で検討した病期の差異



①急性期・回復期の2時点でネットワークの項目に沿ってデータを収集

②ネットワークを生成し、2時点で初学者・熟達者で比較する

図 3.3: 第2実験：病期の変化による初学者群・熟達者群の臨床推論ネットワークの比較検証

だけでなく、模擬患者に提示する情報量および検査異常度の差異によって、臨床推論構造と共通解がどのように異なるのかを検証する。また、提示する患者情報量と検査異常度の影響を検証するため、大腿骨頸部骨折術後の模擬症例として、症例A（検査異常度：低い、情報制限）、症例B（検査異常度：低い）、症例C（検査異常度：高い）の3条件を設定する（図3.4）。また、第3実験では、専門家の合意による正解を仮定せず、参加者集団において列挙頻度が高い上位5位の「問題点と検査」の組み合わせを「共通解」と定義する。各参加者の適合率および再現率を算出し、初学者群と熟達者群の群間比較をMann-WhitneyのU検定を行う。また集約ネットワークとも比較を行い、2.4章で述べた認知バイアスの観点から、推論の収束過程および見落としが生じる条件を明らかにする。

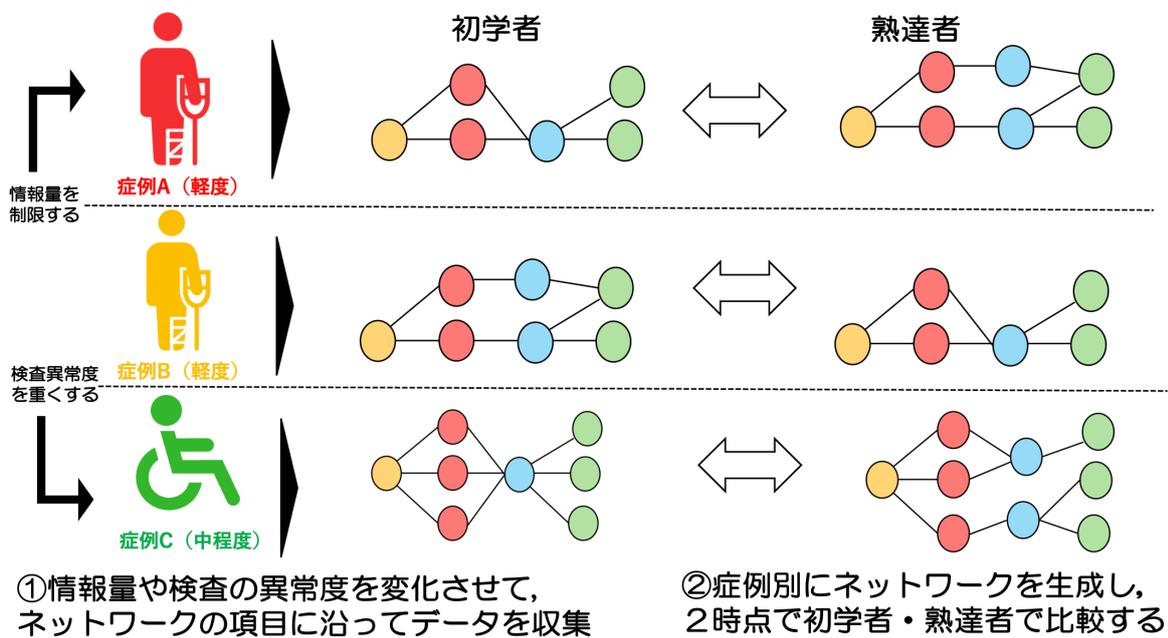


図 3.4: 第3実験：提示情報量および検査異常度の変化が臨床推論ネットワークおよび共通解に与える影響の検証

4 アセスメント文の構造化に関する検証

1.1章で述べたように、理学療法プロセスの「④統合と解釈」は、収集した情報から治療や対処すべき問題点を列挙し、データを統合・解釈して治療すべき問題点を抽出するプロセスである。本研究では、このプロセスにおける思考過程を記述したテキストを「アセスメント文」と定義し、そのアセスメント文を対象とした構造化を行う。

アセスメント文の構造化では、関連研究[33]で課題となった文章間・単語間でのつながりをネットワーク形式で表す。これにより、治療・対処すべき問題点に至るまでの思考過程や対象患者の情報やデータの関係性を示すことができる。この手法を用いることで、従来の文章表現と比較して、臨床推論をより客観的に比較可能な形式へ転換できるかを検証する。

4.1 アセスメント文を構成する項目の設定

本研究では、理学療法士が作成するアセスメント文を対象とする。アセスメント文は自然言語で記述されているため、記述者によって用いる語彙や表現の差異があり形式的に処理することが難しい。そのため、以下の4項目を抽出してからそれらを階層的に繋げることで構造化を図る(図4.1)。階層は第1層から第4層までとし、それぞれの項目に対応させた。

第1層「問題点」

対象者の活動を妨げる因子となり、治療の対象となり得るもの(例:筋力低下、関節可動域制限)。

第2層「対象部位」

上記の問題点が生じている身体部位や関節の運動方向(例:股関節伸展)。

第3層「要因説明」

なぜ上記の問題点が生じたのかを、バイオメカニクスや生理学的観点から説明した文章(例:歩行観察で股関節伸展が不足しているため、歩幅が狭くなっている)。

第4層「根拠データ」

問題点や要因を裏付ける理学療法評価結果やデータ(例:左股関節伸展可動域 -5°)。

表4.1に、図4.1の例文を表形式で構造化した例を示す。

4.2 アセスメント文の構造を表すネットワーク表現

本研究では、4.1章で設定した問題点、対象部位、要因説明、根拠データの関係性を患者ごとに視覚化するため、それらを階層的に繋げたネットワークの生成を試みた。ネットワーク表現を用いることで、表4.1に示すような情報間の関係性を直感的に把握しやすくなり、欠落データの視覚的強調も期待される。

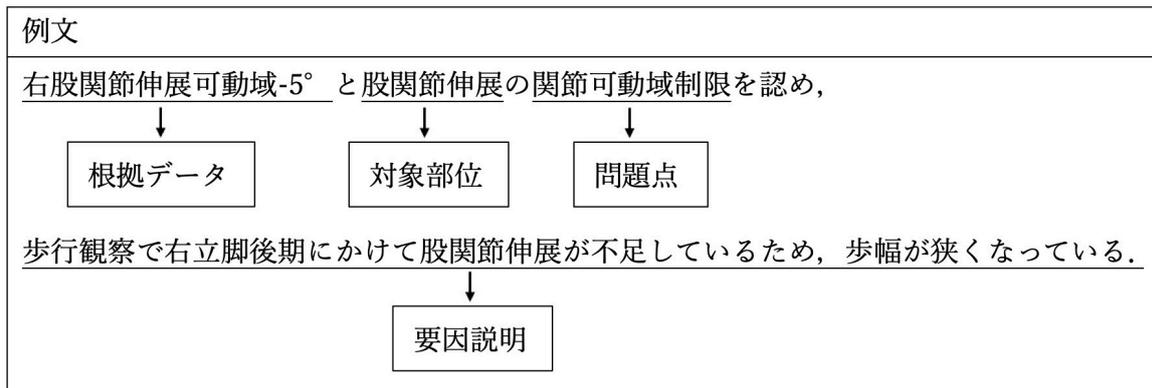


図 4.1: アセスメント文からの各項目の抽出

表 4.1: 各項目の階層

	第1層	第2層	第3層	第4層
意味	問題点	対象部位	要因説明	根拠データ
具体例	関節可動域制限	股関節伸展	歩行観察で右立脚後期にかけて股関節伸展が不足しているため、歩幅が狭くなっている。	股関節伸展筋可動域 -5°

各ノードは特定のデータ項目を示しており、楕円は「問題点」、円形は「対象部位」、四角は「要因説明」、菱形は「根拠データ」をそれぞれ表す。エッジは、これらにある関係性を表現するものであり、関係性の可視化を通じて、データ分析や問題点の理解が容易になることを狙う。要因説明のノードにマウスオーバーすると、該当する要因説明の詳細がポップアップ表示されるように設定した。

ノードとエッジの構造については、中心に「患者」のノードが位置し、外側に向けてノードが階層的に追加される。患者ノードから順に「問題点」→「対象部位」→「要因説明」→「根拠データ」の順にノードが追加され、それぞれの項目が関連づけられていく。すべてのデータ項目が揃っている場合、エッジは黒色で表示される(図 4.2)。一方、データが欠落している場合には、その関係を示すエッジが赤色に変わり、さらにエッジの幅が太く表示されるように設定した。赤いエッジにマウスオーバーすることで、欠落している項目が明示され、欠損情報が視覚的に強調される。例えば、図 4.3 では、股関節伸展(対象部位)の関節可動域制限(問題点)に対する要因説明が欠落しているため、このエッジが赤く太く表示されている。

4.3 検証

本研究では、理学療法士が作成するアセスメント文の構造化が、思考過程やデータの関係性を適切に示すかどうか、またそれらの関係性が臨床推論能力と相関するかを検証する。本稿で対象としたアセスメント文は、吉田ら [33] が実験で収集したデータであり、急性期

病院に勤務する1年目から5年目の理学療法士に、模擬症例（大腿骨頸部骨折症例）をもとに情報分析を実施してもらったもの（計20件）である。参加者の平均経験年数は2.45年、アセスメント文の平均文字数は652.5文字であった。アセスメント文に関しては盲検化され、理学療法養成校の大学教員によるパフォーマンス評価が行われた。評価は臨床推論のルーブリック [24] に基づいて15点満点で行われ、その平均点は5.9点であった。なお、実験で用いた模擬症例に関する「問題点」の正答は、大腿骨頸部骨折ガイドライン [25][26] に沿って理学療法養成校の大学教員（本稿第4著者）と認定理学療法士（本稿第3著者）の合議により設定された。

4.4 検証結果

アセスメント文の構造化の定量評価を行うために①経験年数、②パフォーマンス評価点数、③アセスメント文に挙げられた問題点の数、④設定した正答に対する問題点の再現率、⑤設定した正答に対する問題点の適合率、⑥再現率と適合率の調和平均であるF値を求めた。なお、今回はアセスメント文の構造化（e.g., 図4.4）は人手で行った。

経験年数を基準とした相関分析を行ったところ、経験年数と問題点の数に中程度の負の相関（ $r = -0.48$ ）、経験年数と適合率に、弱い正の相関（ $r = 0.37$ ）、経験年数と再現率に弱い負の相関（ $r = -0.34$ ）が各々認められた。

高いパフォーマンス評価のアセスメント文の構造化（図4.4）とそれに基づいたネットワークを示す（図4.5）。図4.4のように、問題点に対して、その問題が発生している対象部位、要因説明、根拠データが明示されている場合、構造化が適切に行われていると判断した。

低いパフォーマンス評価の場合、アセスメント文の中でも構造化されていない文章が多く、問題点に対する説明が曖昧であり、要因説明や根拠データの記載が不足している傾向が見られた。低いパフォーマンス評価のアセスメント文の例を図4.6、ネットワークを図4.7に示す。図4.6の文章のように、問題点に対して対象部位の記載があったとしても、要因説明が不足しており、具体的な理学療法評価結果などの根拠データの記載がなかった。

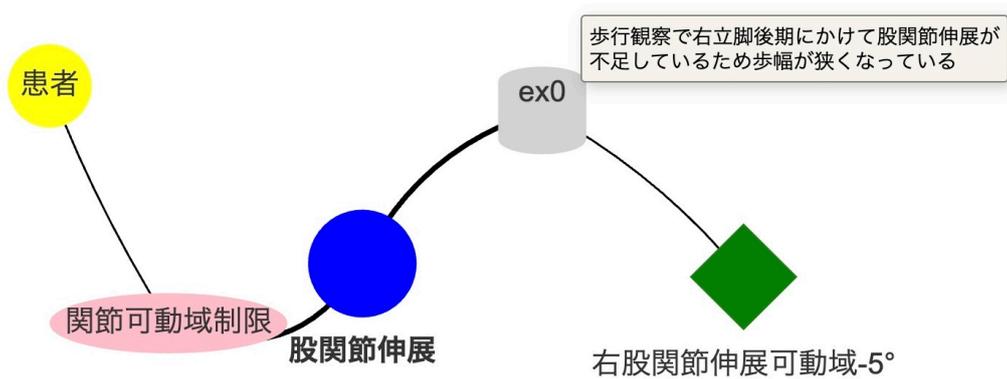


図 4.2: 情報の欠落のない文章

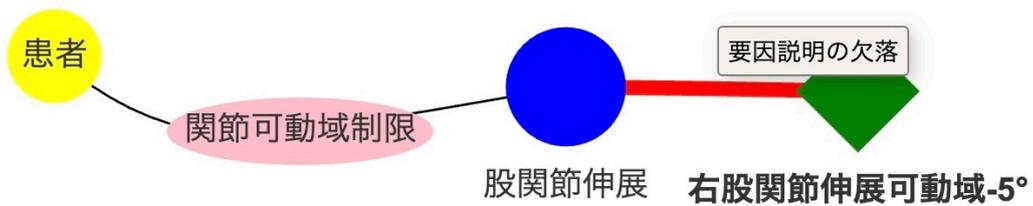


図 4.3: 要因説明が欠落している文章. アセスメント文には右股関節伸展可動域が -5° であることが示されているものの, それが股関節伸展の説明の根拠であることが文章として明示されていない.

パフォーマンス評価: 11/15点 (平均)			
歩行動作において、①右荷重応答期から立脚中期においてデュシャンヌ様の跛行が生じている要因としては、検査より股関節外転筋 MMT2, 伸展筋 MMT2が挙げられる。デュシャンヌ様跛行では主に中殿筋の筋力低下が生じることにより生じやすいとされており、本症例においても股関節伸展, 外転の筋力が MMTで2レベルであることから、デュシャンヌ様跛行を呈していると考えられる。			
第1層「問題点」	第2層「対象部位」	第3層「要因説明」	第4層「根拠データ」
筋力低下	股関節外転	右荷重応答期から立脚中期においてデュシャンヌ様の跛行が生じている	股関節外転筋 MMT 2
	股関節伸展	右荷重応答期から立脚中期においてデュシャンヌ様の跛行が生じている	股関節伸展筋 MMT 2

図 4.4: 高いパフォーマンス評価のアセスメント文の構造化 (抜粋)

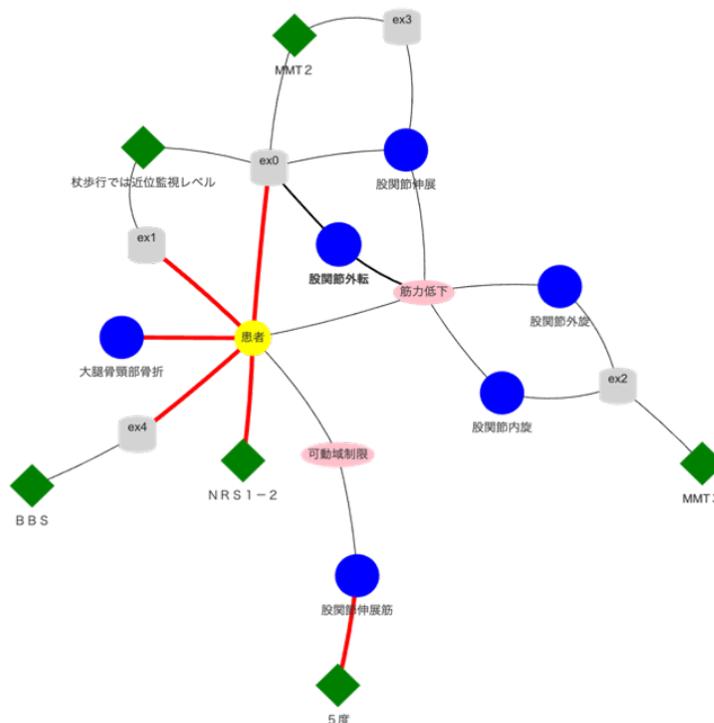


図 4.5: 高いパフォーマンス評価のアセスメント文のネットワーク構造

パフォーマンス評価: 1.5/15点 (平均)			
歩行においては右荷重応答期・立脚中期にかけてのデュシャンヌ様の跛行, 立脚後期での股伸展減少, 右立脚期短縮を認めた。理学療法評価より, 右股関節伸展筋力低下・外転筋力低下が問題点としてあげられる。右股関節自動伸展可動域および筋力低下による最大可動域まで達していないと考えられる。			
第1層「問題点」	第2層「対象部位」	第3層「要因説明」	第4層「根拠データ」
歩行		右荷重応答期・立脚中期にかけてのデュシャンヌ様の跛行, 立脚後期での股伸展減少, 右立脚期短縮を認めた。	
筋力低下	右股関節伸展		
筋力低下	右股関節外転		

図 4.6: 低いパフォーマンス評価のアセスメント文の構造化 (抜粋)

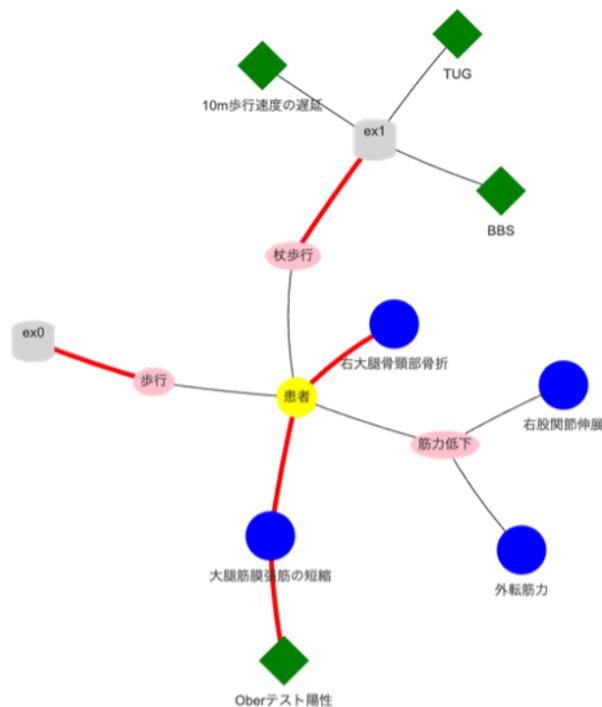


図 4.7: 低いパフォーマンス評価のアセスメント文のネットワーク構造

4.5 考察

相関分析の結果、経験年数と列挙された問題点の数の間に中程度の負の相関、経験年数と適合率の間に弱い正の相関が各々認められたことから、経験年数が増加するにつれて指摘する問題点が減少するが、挙げられた問題点は正確なものである傾向があることが示唆された。一方、経験年数の浅い理学療法士は、多くの問題点を指摘する傾向が見られ、的確に問題点を特定するというよりも、幅広く指摘する傾向が確認された。また、経験年数と再現率に、弱い負の相関が認められ、経験を重ねた理学療法士でも、全ての問題を的確に列挙できていない可能性が示唆された。この結果は、見落としのリスクが増加している可能性を示唆している。経験年数の増加に伴い扱ったことのある症例数が増えるため、過去の症例経験に照らして「おそらくこれだろう」という直感が働くために問題点を指摘する能力が向上するため、余分な機能面の指摘が減り、適合率が向上したと考える。しかし、その直感的な判断が、対象者の観察や仮説検証の過程を疎かにし、見落としを招いた懸念がある。

経験年数とF値の間に相関が認められなかった($r = -0.07$)ことから、臨床経験による問題点の特定能力の向上という正の側面と、馴化に伴う見落としという負の側面が相殺され、結果として「経験年数が増えたからといって、必ずしも的確な臨床推論ができていないわけではない」という可能性が示唆される。これは、文献 [32] で指摘された「経験年数とリスク列挙数の間に相関がない」という結果とも符合する。

また、ネットワークで可視化することによって、臨床推論を客観的に比較することが可能となった。図 4.5 と図 4.7 を比較すると、図 4.7 (低いパフォーマンス評価) のネットワークは、図 4.5 (高いパフォーマンス評価) のネットワークに比べてノード数・エッジ数が少ない。以上より、本研究で提案した臨床推論の構造化およびネットワーク可視化により、テキストのままでは比較が難しい推論内容を、比較可能な形式へ変換できる可能性が示された。

4.6 本章のリミテーション

本研究のアセスメント文の構造化は人手で行ったため、文章の書き方や文章量によってネットワークが変化した。例えば、根拠データを省略している文章は今回は根拠の欠落として短いネットワークとなる。例文としては「歩行においては右荷重応答期・立脚中期にかけてのデュシェンヌ様の跛行、立脚後期での股関節伸展減少、右立脚期短縮を認めた」という文章では、問題点は「歩行」、要因説明は「右荷重応答期・立脚中期にかけてのデュシェンヌ様の跛行、立脚後期での股関節伸展減少、右立脚期短縮を認めた」となり、対象部位と根拠データの欠落が示されるネットワークになる。しかし、「デュシェンヌ様の跛行(以下、デュシェンヌ徴候)」は理学療法学では股関節外転筋力の低下によって引き起こされる歩容の一種であり、理学療法養成校で学ぶ基本的な理学療法で使用される専門用語である。そのため、今後の課題としては、アセスメント文を構造化する上で「デュシェンヌ徴候」のような理学療法に関する用語のリストを作り、構造化することで、文章の丁寧さに左右されない構造化手法を実現する必要性が挙げられる。

また、本研究におけるアセスメント文の構造化は単一分析者(著者)によって実施したため、抽出結果に分析者の解釈が影響しえる点、ならびに構造化の再現性を十分に検証でき

ていない点がリミテーションとして挙げられる。今後は、複数分析者による独立した抽出作業を行い、構造の一致率を算出することで、構造化手法の再現性を検証する必要がある。

なお、本研究の立ち位置として、臨床推論の正誤性を判定することよりも、推論過程を構造化し、理学療法士間で比較可能な形式へ変換することを主目的とした。したがって、本研究で得られたネットワークは「正解」を表すものではなく、各参加者がどのような情報を根拠として、どのように問題点へ結び付けたかという推論過程の表出として解釈されるべきである。

4.7 本章の結論

本章では、理学療法プロセスにおけるアセスメント文を構造化し、ネットワークを生成することで、思考過程を可視化する手法を検討した。構造化の結果、経験年数の増加に伴い、必ずしも適切ではない問題点の選択数は減少し、より正確な問題点が挙げられる傾向が見られた。しかし、全ての問題を的確に特定できているわけではないことが示唆された。また、経験年数が少ない理学療法士では、相対的により多くの問題点を指摘するが、その正確さに欠ける傾向が確認された。これらの結果から、有用であると考えられるシステム構築のために初学者理学療法士にわかりやすいネットワーク表現を用い、より正確な問題点へ導き、適切な臨床推論が行えるのについて検証を行う必要があると考えられる。

また、経験年数とF値の間に相関が認められなかったため、経験年数が増えたからといって必ずしも的確な臨床推論ができるわけではないことが分かった。このことは、臨床推論能力は現場経験によって熟達によってのみ育成されるのではないことを示唆しており、経験に伴う直感力の向上に加えて、見落としリスクの支援などにより適切な能力の育成の可能性が示唆された。

5 病期の差異による初学者群・熟達者群の臨床推論構造の比較検証

5.1 ネットワークの項目の設定

提案手法1でリミテーションであった、文章の丁寧さに左右されない構造化手法を実現するために、本章ではアセスメント文等の記述文を用いるのではなく、構造化した項目のみ選択式でデータを収集し、ネットワークで表す。本章では、以下の3層でネットワーク生成する。ネットワークはPythonライブラリであるNetworkXを用いて、「患者→問題点→身体部位→検査」の有向グラフを作成した(図5.1)。本章では文章を扱わないため、前章の第3層目に当てはまる「要因説明」を除いた項目で構成した。ネットワークの構造の例を挙げる。例えば、膝関節伸展のMMT、圧痛の有無を検査を行い、大腿四頭筋に痛みや筋力低下を認めた対象者がいるとする。この場合、ネットワークは、問題点に「痛みに感覚」、「筋力の機能」、対象部位に「大腿四頭筋」、データに「圧痛」「膝関節伸展MMT」が構成される(図5.2)。このようにネットワークで表すことで、図5.3で表すように、問題点と対象部位、データとの関係を表すアセスメント文と比較しても、視覚的にどこに着目しているかがわかりやすくなる。

第1層「心身機能の問題点」

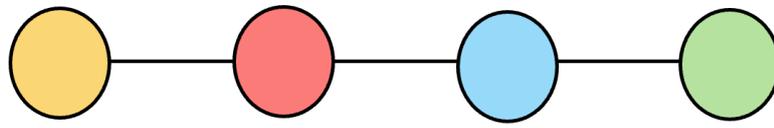
対象者に当てはまる心身機能の問題点。心身機能はICFで定められているコードを用いる。(例：筋力の機能=ICFコードb730, 痛みの感覚=ICFコードb280)。

第2層「対象部位」

上記の問題点が生じている身体部位や関節の運動方向(例：股関節外転, 大腿直筋)。

第3層「検査データ」

問題点や要因を裏付ける理学療法評価結果やデータ(例：MMT股関節外転, ROM-t股関節伸展)。



患者 → ①心身機能の問題点 → ②対象部位 → ③データ

図 5.1: ネットワークの構造

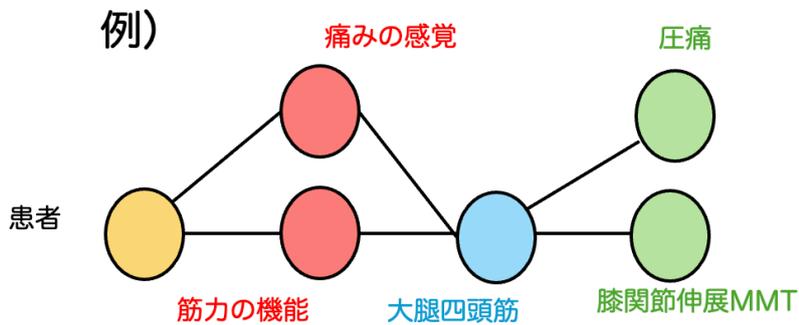


図 5.2: ネットワークの例

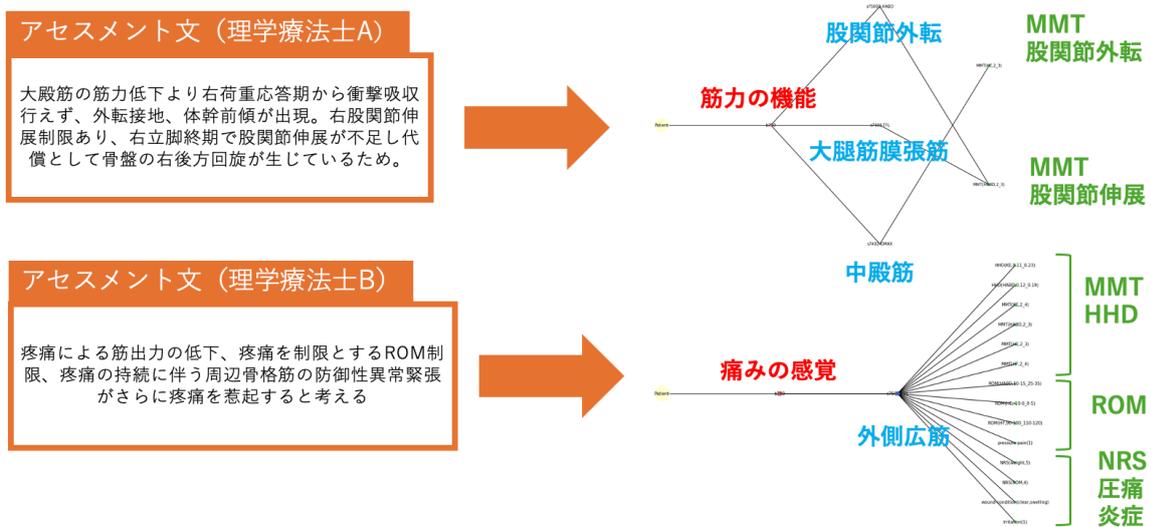


図 5.3: アセスメント文とネットワークの比較

5.2 実験目的

本実験の目的は3つある。第1の目的は、文章量や文章表現の丁寧さに左右されない形で臨床推論を構造化することである。4.6章で述べたリミテーションに対応するため、本実験ではアセスメント文から臨床推論を抽出するのではなく、ネットワークの項目に対応した質問項目からなるアンケート調査を実施する。ネットワークの項目に沿って回答を求めることで、文章では省略され得る検査項目や判断過程を明示化できると考える。これは、1.5章で述べた臨床推論教育の課題である「なぜその治療を選択したのか」という判断過程を可視化することができると思う。

第2の目的は、単一時点における経験年数による比較にとどまらず、病期（急性期→回復期）の差異が臨床推論がどのように影響を与えるのかを明らかにすることである。1.3章で述べたように、理学療法士には急性期医療から回復期医療にかけての関与の必要性が高まっている。病期に応じて、限られた時間と情報の中で問題を同定し、優先順位を付け、治療の根拠と見通しを言語化する臨床推論がより強く求められる。そのため本実験では、同一の模擬症例を用い、手術直後の急性期時点と、手術から日数が経過した回復期時点の2条件の情報を提示し、それぞれアンケート調査を実施する。急性期から回復期にかけての回答を比較することで、理学療法士が着目する検査や問題点がどのように変化するか、ならびに臨床推論構造の特徴を検討する。

第3の目的は、経験年数によって臨床推論に差異がみられるかを、推論構造に基づいて比較することである。4章の結果では、経験年数が高いほど列挙された問題点の数が増加する傾向が示された。このことから、本実験においても経験年数が高いほど臨床推論の幅が大きくなると仮説する。また、着目する視点や臨床推論が経験年数によって変化するかを明らかにするため、初学者群と熟達者群に分けて群間比較を行う。

5.3 実験方法

経験1～2年の初学者群(10名)と経験6～13年の熟達者群(10名)に対して、2つの模擬症例を提示し、アンケート調査を行った。模擬症例は、大腿骨転子部骨折術後7日目(POD7)と30日目(POD30)の2つデータをGoogleスプレッドシートで提示し、必要な検査、起こりうる心身機能の問題点(最大5個まで)と対象部位(個数制限なし)、問題点を導くために必要な検査(個数制限なし)をGoogle Formを用いて選択させた。

模擬症例は、大腿骨転子部骨折術後のクリニカルパス[20]と比較すると遅れているが、長期間の入院・リハビリテーションを行うことで杖歩行を目指せるレベルに設定した。POD7の時点では、トイレまでは看護師の見守りで歩行器歩行で歩けるが、長距離は車椅子介助で移動するレベルの症例像にした。POD30の時点では、病棟内は一人でトイレに行けており、自宅復帰に向けて杖歩行を理学療法で練習中である症例像にした。また検査の数は、社会的情報(住環境、受傷前の生活など)は約20個、医学的情報(血液データ、手術記録)は約30個、理学療法検査(筋力検査、関節可動域検査、歩行テスト、日常生活動作)は約150個をスプレッドシートで提示した。

前述した手法を用いて「患者→問題点→身体部位→検査」のネットワークを作成し、ノード数、エッジ数を調査した。ノード数は参加者が回答した問題点、身体部位、検査の総数



図 5.4: 模擬症例 (POD7 → POD30)

を示し、エッジ数はそれらを繋ぐ実線の本数を示す。ノード数は「臨床推論の広がり」の指標とし、エッジ数は「問題点・部位・検査の関係付け」の指標とする。

5.4 検証と結果

5.4.1 検証1：記述統計と代表例のネットワークの比較

検証1では、POD7とPOD30の初学者群・熟達者群のノード数とエッジ数による比較を行う。まずは、初学者・熟達者のノード数、エッジ数の平均と標準偏差を示す。また特徴的であったPOD7・POD30の初学者・熟達者のネットワークを生成し、視覚的に比較を行う。

POD7・POD30の初学者群・熟達者群の記述統計の結果を表5.1で表す。POD7・POD30ともに熟達者の方がノード数・エッジ数ともに多い傾向であった。また特徴的だったネットワークを図5.5, 図5.6と図5.7, 図5.8で表す。まずは理学療法士9年目の熟達者のネットワークに着目する。POD7では、股関節周囲の「筋力の機能」や「関節の可動性」「痛みの感覚」などの主に理学療法の対象となる問題点のみならず、「血液の機能」のリスク管理面での問題点を挙げてきていた(図5.5)。また、検査も徒手筋力テスト(Manual Muscle Test: 以下MMT)や関節可動域測定(Range Of Motion Test: 以下ROM-t)のみならず、医学的情報に含まれる血液検査のヘモグロビン値や医師記録、手術中の出血量などの貧血に関する検査を挙げていた。POD30では、「筋力の機能」や「痛みの感覚」のみならず、「筋の持久性」などの筋肉に対する問題点が多く挙げられていた(図5.6)。検査ではMMTや痛みの検査のみならず、全身持久力を測定するテストの6分間歩行や日常生活動作(Activities of Daily Living: 以下ADL)など自宅復帰に向けた検査が挙げられていた。一方で、理学療法士1年目のネットワークに着目すると、POD7・POD30ともに問題点は「筋力の機能」しかあげておらず、検査も歩行観察やMMT, ROMと病期に応じた大きな差異は見られなかった(図5.7, 図5.8)。

これらの結果から、臨床推論をネットワークで可視化することにより、客観的に比較できた。病期による思考の変化だけでなく、理学療法士の思考の違いを観察できた。推論過

時点	群	ノード数			エッジ数		
		平均 ±SD	中央値	範囲 (min-max)	平均 ±SD	中央値	範囲 (min-max)
POD7	初学者	25.8 ± 17.5	23	6-51	106.4 ± 115.2	59.5	5-307
	熟達者	34.5 ± 20.8	33	10-77	145.5 ± 174.5	79.5	14-477
POD30	初学者	15.0 ± 7.8	14	5-27	38.8 ± 37.7	19.0	4-100
	熟達者	18.4 ± 14.1	15	7-54	75.6 ± 113.4	28.5	7-373

表 5.1: POD7 および POD30 におけるノード数・エッジ数の記述統計 (平均 ± 標準偏差, 中央値, 範囲)

程を可視化する手法は, 教育支援に有効な客観的評価であると考えられる。

術後7日目 ノード数=40, エッジ数=110

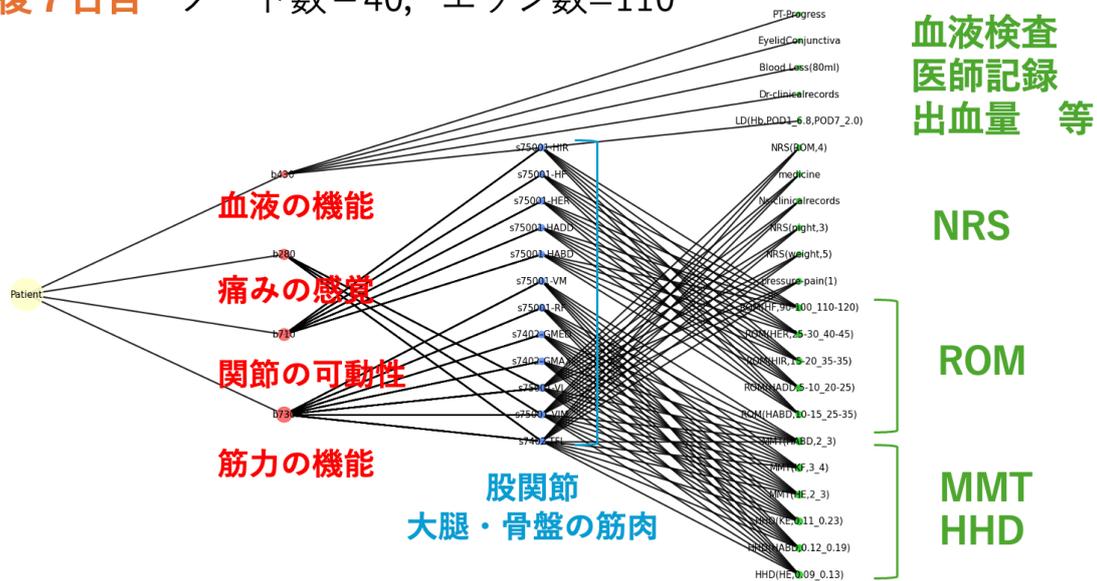


図 5.5: 代表例：熟達者（9年目）のPOD7のネットワーク

術後30日目 ノード数=22, エッジ数=56

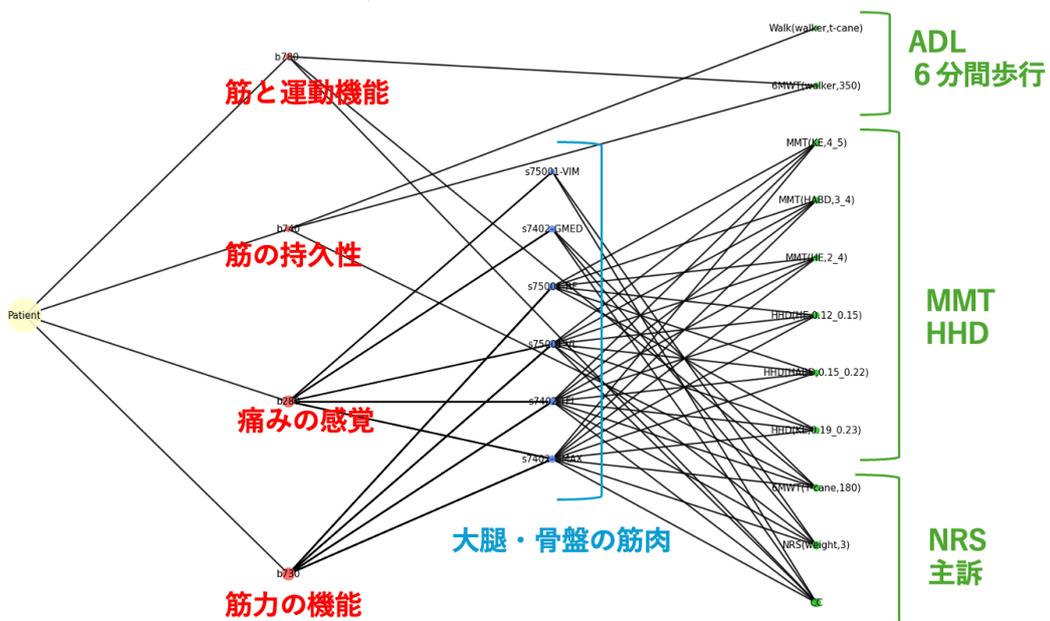


図 5.6: 代表例：熟達者（9年目）のPOD30のネットワーク

術後7日目 ノード数=6, エッジ数=5

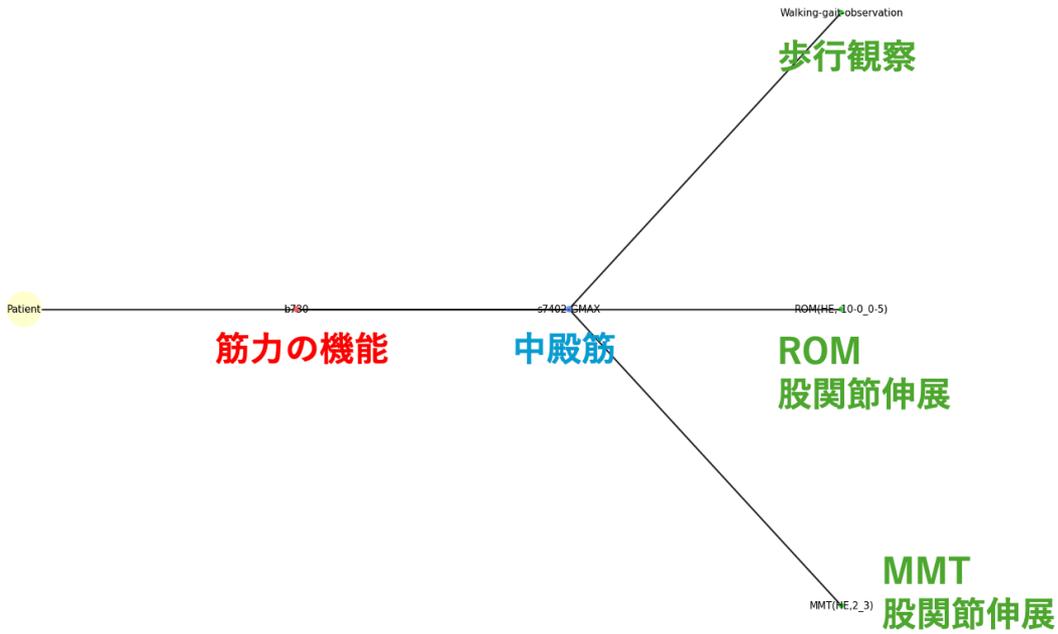


図 5.7: 代表例：初学者（1年目）のPOD7のネットワーク

術後30日目 ノード数=9, エッジ数=9

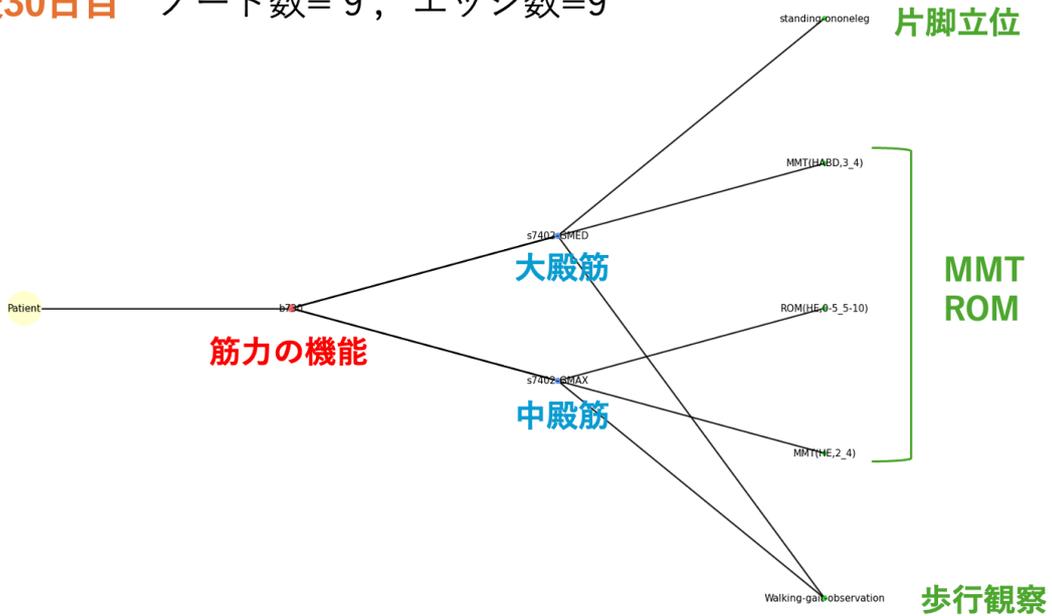


図 5.8: 代表例：初学者（1年目）のPOD30のネットワーク

5.4.2 検証 2：初学者群・熟達者群の群間比較

検証 2 では、検証 1 の結果を踏まえて、ノード数・エッジ数ともに熟達者の方が有意に多いのではないかと仮説する。つまり、熟達者の方が臨床推論の幅が広く、問題点・対象部位・検査の関係付けが多くできているのではないかと考える。そのため、本検証では、POD7 におけるノード数・エッジ数の初学者群・熟達者群、POD30 におけるノード数・エッジ数の初学者群・熟達者群の Mann-Whitney U 検定で群間比較を行う。

Mann-Whitney U 検定の結果、表 5.2 に示す。初学者群と熟達者群の POD7・POD30 でのノード数・エッジ数には有意な差は得られなかった（表 5.2）。この結果の考察としては、表 5.1 で表したように、ノード数・エッジ数ともに範囲（min-max）が広く、個人差が大きかった。このようなばらつきの大きさは群間差の検出力を低下させる可能性がある。初学者群・熟達者群ともに問題点・身体部位・検査を絞り込まず、検査結果がカットオフ値より低いものはたくさん選んでいる参加者が散見された（図 5.9）。また反対に、問題点・身体部位・検査が極端に少ない参加者もいた（図 5.10）。そのため、本検証では群間差が統計的に有意として検出されなかった可能性がある。

5.4.3 検証 3：術後経過による変化（POD7 と POD30 の比較）

検証 3 では、急性期の POD7 から回復期の POD30 にかけてノード数、エッジ数ともに有意に減るのかを検証する。仮説は、熟達者の方が POD7 から POD30 にかけて有意にノード数・エッジ数ともに減り、問題点や必要な検査が絞り込まれるのではないかと考える。検証は、参加者全員と初学者・熟達者の 3 群内で対応のある Wilcoxon 符号付順位検定を行う。

対応のある Wilcoxon 符号付順位検定の結果を、表 5.3 に示す。参加者全員のノード数・エッジ数ともに POD7 から POD30 にかけて有意に減少した。また初学者群のノード数のみ、POD7 から POD30 にかけて有意に減少した。熟達者もノード数・エッジ数ともに減少傾向であるが、有意な差は示さなかった。

指標	時点	初学者（中央値）	熟達者（中央値）	U	p	(r)
ノード数	POD7	23.0	33.0	38.5	.406	(-0.23)
	POD30	14.0	15.0	47.0	.849	(-0.06)
エッジ数	POD7	56.0	65.0	43.0	.623	(-0.14)
	POD30	19.0	28.5	41.5	.544	(-0.17)

表 5.2: 初学者群と熟達者群の群間比較（Mann–Whitney U 検定）

指標	群	POD7	POD30	n	p	r
ノード数	参加者全員	27.5	15.0	18	0.004	-0.772
	初学者	23.0	14.0	8	0.042	-0.833
	熟達者	33.0	15.0	10	0.053	-0.709
エッジ数	参加者全員	59.5	20.0	19	0.009	-0.689
	初学者	59.5	19.0	9	0.109	-0.622
	熟達者	79.5	28.5	10	0.083	-0.636

表 5.3: POD7 から POD30 への変化：Wilcoxon 符号付順位検定（対応あり）

5.4.4 検証4：集約ネットワークによる比較

検証4では、POD7とPOD30で初学者群と熟達者群の集約ネットワークの比較を行う。集約ネットワークとは、初学者群・熟達者群に属する参加者の個別ネットワークを統合して得られるネットワークである。集約ネットワークで比較する点としては2点ある。1つ目は、POD7またはPOD30で初学者群・熟達者群それぞれどの問題点や検査に着目しているのかを比較する。2つ目は、POD7からPOD30にかけて着目している点がどのように変化するかを比較する。

各群の参加者すべてのエッジ・ノードを統合すると、ノード数・エッジ数が膨大になり、集約ネットワークが示す特徴を掴むのが困難となる。そのため、エッジ表現を変更する。「問題点」「身体部位」「検査」の3つ組み合わせを参加者が5人以上挙げていると赤エッジ、参加者が4人以上挙げていると青エッジ、参加者が3人以上挙げている場合は緑エッジとした。また参加者が2人以下の組み合わせは非表示とした。

POD7の熟達者群の集約ネットワークを図5.11、初学者群の集約ネットワークを図5.12に各々示す。まずは、熟達者群の集約ネットワークに着目する。問題点には「筋力の機能」「関節の可動性」「痛みの感覚」が挙げられていた。「筋力の機能」「関節の可動性」の検査には、MMTやハンドヘルドダイナモメーター（Hand Held Dynamometer：以下HHD）の筋力検査やROMなどが挙げられていた。「痛みの感覚」では、圧痛や痛みの主観的検査（Numerical Rating Scale：以下NRS）へ赤エッジが伸びていた。一方で、初学者群の集約ネットワークでは、問題点が「筋力機能」のみにしか挙げられておらず、検査もMMTやHHDなどの筋力に関する検査が挙げられていた。

POD30の熟達者群の集約ネットワークを図5.13、初学者群の集約ネットワークを図5.14に示す。POD30の模擬症例は歩行器歩行で病棟内を歩行できているが、杖歩行が未獲得であり、階段などの段差練習中に痛みがあるという症例であった。熟達者群の集約ネットワークでは、「筋力の機能」のみならず「痛みの感覚」を問題点としてあげている参加者が多かった。検査はMMT・HHDの筋力検査の他に荷重時のNRSをあげており、またPOD7には見られなかった10m歩行（歩行の実用性・効率性や転倒リスクの評価）の歩行のパフォーマンス評価が見られた。一方で、初学者群は問題点はPOD7と同じく「筋力の機能」をあげている参加者が多く、検査はMMT・歩行観察・片脚立位と筋力や歩行に着目した検査が多かった。

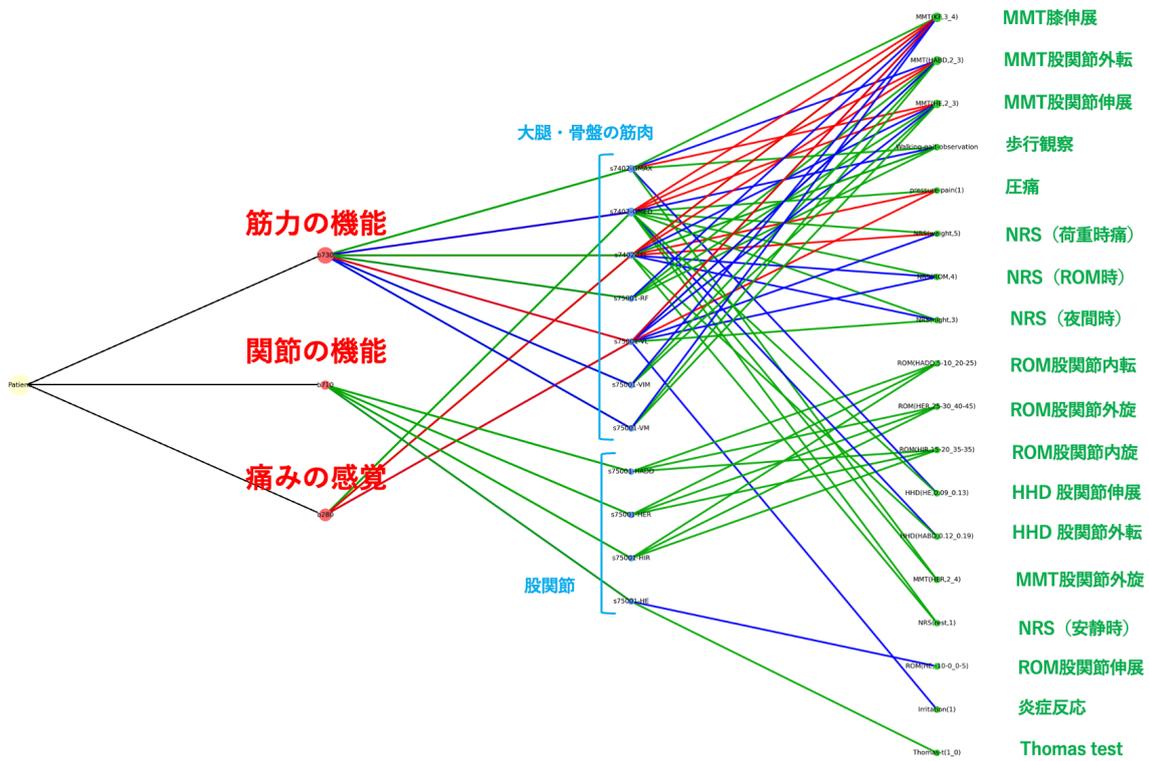


図 5.11: POD7の熟達者の集約ネットワーク

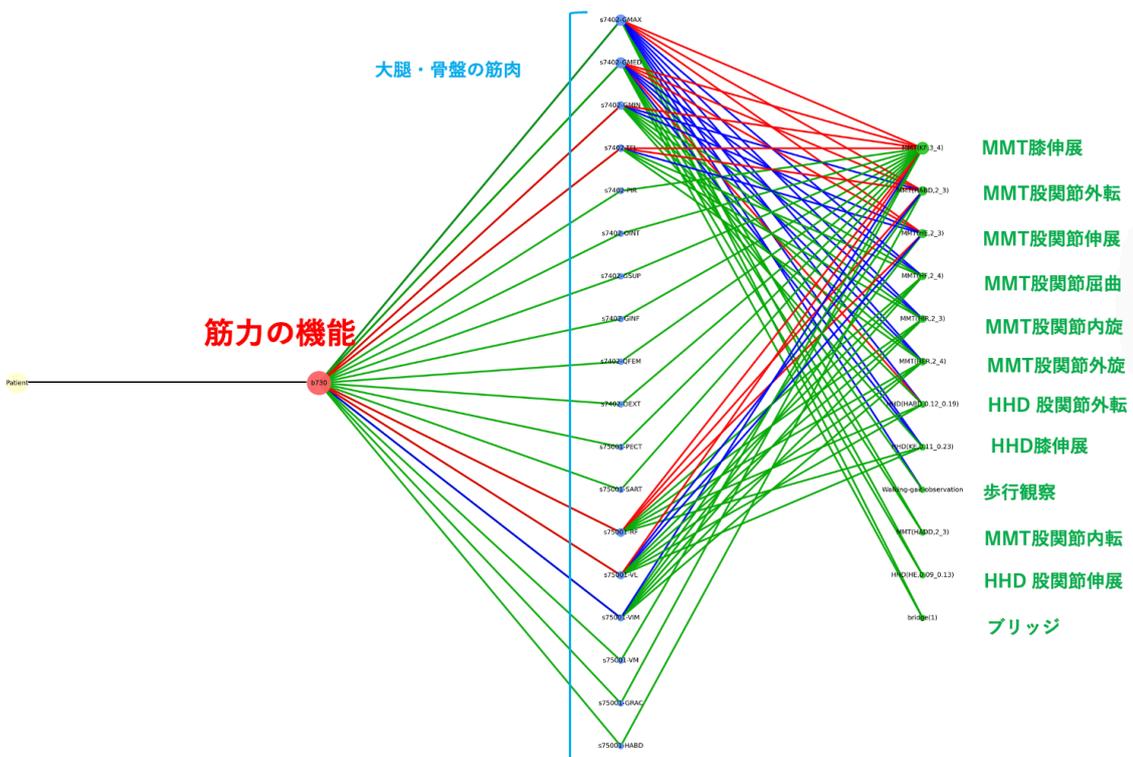


図 5.12: POD7の初学者の集約ネットワーク

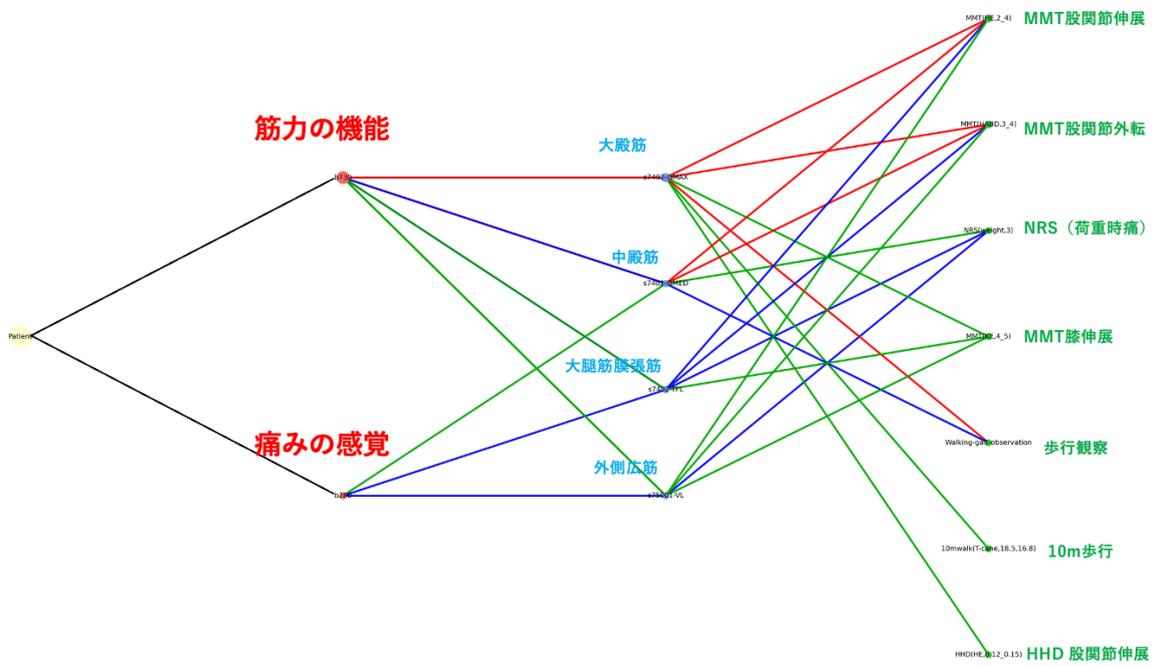


図 5.13: POD30 の熟達者の集約ネットワーク

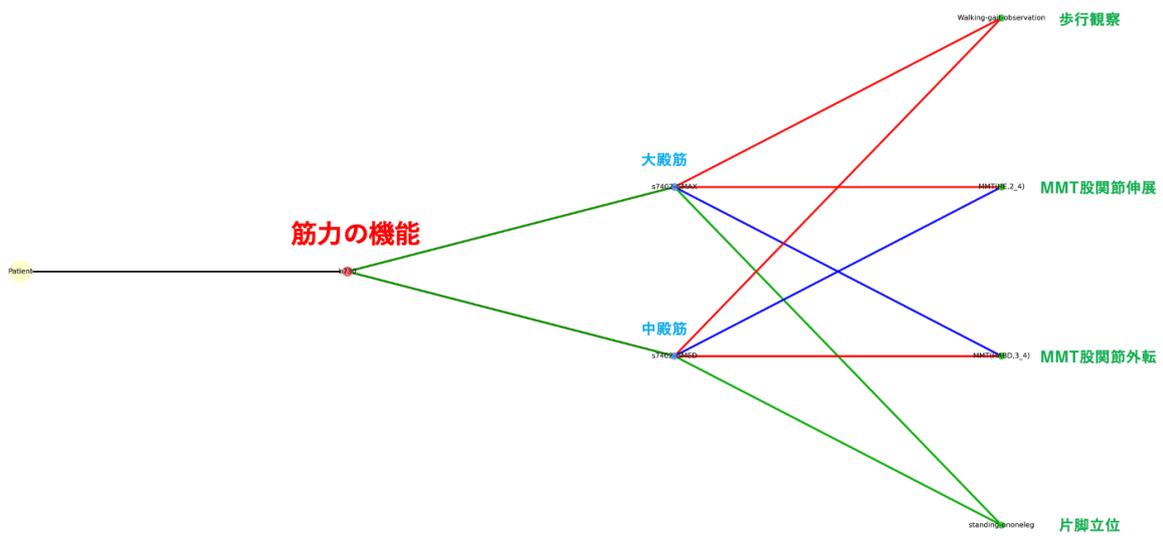


図 5.14: POD30 の初学者の集約ネットワーク

5.5 考察

検証2では、「問題点・身体部位・検査」から構成されるネットワークのノード数およびエッジ数について、初学者群と熟達者群の間に有意差は認められなかった。検証1の結果を踏まえると、両群ともにノード数・エッジ数のばらつき(max-min)が大きかった。すなわち、わずかな異常所見から多くの「問題点・身体部位・検査」を列挙する参加者がいる一方で、検査結果を踏まえて問題点や身体部位を慎重に絞り込む参加者もいた。このような参加者間の違いが大きかったため、検証2の群間比較では差が検出されなかった可能性がある。

次に、検証3および検証4の結果から、初学者群ではPOD7からPOD30にかけてノード数およびエッジ数が有意に減少した。検証4の集約ネットワークの所見と併せて考察すると、熟達者群はPOD7の時点で「筋力の機能」「関節の可動性の機能」「痛みの感覚」など複数の問題点を挙げ、筋力検査、関節可動域検査、痛みの評価など多様な検査を列挙する傾向がみられた。さらにPOD30においても、「筋力の機能」「痛みの感覚」およびそれらに関連する検査が継続して挙げられる参加者が多かった。一方、初学者群はPOD7・POD30を通じて主に「筋力の機能」を中心に指摘する傾向がみられ、POD30ではMMTの測定項目数が減少していた。このことから、初学者群はもともと着目範囲が限定的であり、術後経過の進行に伴ってさらに選択が収束した結果、ノード数・エッジ数の有意な減少として表れたと考えられる。

以上より、本章では「初学者群と熟達者群の臨床推論の幅に大きな差はみられない」こと、および「初学者群では術後経過に伴い推論が収束しやすい」ことが示唆された。しかし、これらの結果は、4章の考察で述べた「経験年数を重ねるほど問題点が減少する」という所見とは整合しない。この不整合は、データ取得方法の違いに起因すると考えられる。6章ではアセスメント文から問題点を抽出したため、記述量や表現方法により抽出可能な問題点数が限定される。一方、本実験では文章を介さず、問題点・身体部位・検査を選択式で回答させ、選択数に上限を設けなかった。そのため、文章には顕在化されにくい着目点や検査候補が、本章の実験方法によって表出した可能性がある。すなわち、熟達者は文章化の段階では正確性の高い内容のみを記述する一方で、選択式の課題では背景にある探索的な着目点や候補をより広く提示すると推察される。

5.6 本章のリミテーション

本章のリミテーションは2つある。第1に、理学療法士の臨床推論の再現性を十分に明らかにできなかった点である。本章では模擬症例を用いて実験を実施したが、初学者群・熟達者群のいずれにおいても、理学療法士が着目する検査や列挙する問題点には大きな個人差が認められた。この差異は、臨床で日常的に担当する患者層や病期、ならびに経験してきた症例の違いに影響を受けた可能性がある。特にPOD7の熟達者群では、痛みや炎症反応などの「リスク管理」を重視する者や、関節機能に重点を置く者がみられた一方、初学者群では筋力に重きを置く傾向があった。これは、1.6章で述べた課題、すなわち同じ国家資格を有する理学療法士間でも推論の再現性が必ずしも担保されていないという状況と整合する。したがって次章では、模擬患者の検査異常度・重症度を段階化し、経過が順調な

症例や一般的な経過に比べて回復が遅い症例を設定することで、どのような条件で推論の特徴や偏りが生じやすいかを検討する。また、参加者間での共有されやすく重要である解として「共通解」を設定し、それに対する再現率および適合率を算出することで、再現性を定量的に評価する。

第2に、本章では初学者群・熟達者群のネットワーク指標（ノード数・エッジ数）および集約ネットワークにおけるパスパターンに着目したものの、推論の妥当性を定義できていなかった点である。そのため、各参加者が選択した問題点や検査が臨床的に妥当であるかを直接評価することは困難であった。この点を補うため、次章では問題点と検査の「共通解」を算出し、それに対する適合率および再現率を用いて各参加者の思考過程を評価する。

6 提示情報量と検査異常度の差異が臨床推論ネットワークおよび共通解に及ぼす影響の検証

6.1 「共通解」に関する研究と定義

5.6章で述べたリミテーションを解決するため、「共通解」について関する研究や定義づけを行なっていく。医療や教育分野では、正解が明確ではない場合に複数人の意見から「コンセンサス（合意解）」を定義し、それを基準として個々の回答を評価する方法が用いられている。例えば、医学教育の研究では、根拠が不十分な領域では専門家の判断を合意することでガイドラインや評価基準を作成する [7]。Consensus methods を用い、専門家の意見を反復的に調整し「グループ合意」を得る方法が一般的であり、これらは「専門家の判断を厳密に統合」し、賛同度を測る手段として利用されている。Warby らは、複数の熟練評価者によるスコアリングから得たグループ合意をゴールドスタンダードとして用いた。また、各評価者の性能は、当該評価者を除外して作成した leave-one-out 合意を参照基準とし、適合率および再現率により評価した。これは、合意を共通解とみなし、個々の判断が共通解とどの程度一致するかを定量化する枠組みである。

また国際的な報告ガイドラインである STARD2015 (Standards for Reporting Diagnostic Accuracy Studies) によれば、参照基準（リファレンススタンダード）は「ターゲットとなる状態（疾患など）の有無を確立するために利用可能な最良の方法」と定義されている。理想的には、エラーを含まない「ゴールドスタンダード」が望ましいが、臨床の現場において 100% の精度を持つ検査法は稀である。そのため、STARD では「ゴールドスタンダード」という用語よりも、実用上最良の方法という意味を込めて「参照基準」という用語の使用を推奨している。 [3]。

以上の先行研究を踏まえ、本研究における「共通解」を以下のように定義する。本研究では、認定理学療法士などの専門家による合意に基づくゴールドスタンダードや参照基準を仮定するのではなく、参加者間において頻度が高く、共有されていた「問題点（心身機能・活動・参加）と検査」の組み合わせを「共通解」と定義する。この共通解は、正解そのものを示すものではなく、臨床推論において参加者の判断がどの程度収束しているか、あるいは共有されやすい重要な組み合わせが何であることを示す指標である。本研究では、この共通解を用いて各参加者の臨床推論が集団における判断の共有性とどの程度一致しているかを、適合率および再現率を用いて評価する。

6.2 ネットワークの項目の設定

今回の検証では、ICF の構成要素となる「心身機能」だけでなく、「活動」と「参加」もネットワークの項目として設定する。「活動」には歩くや階段を登るなどの身体活動にすることが当てはまり、「参加」には履き物を履くなどのセルフケアや趣味、買い物などが当てはまる。今回は問題点と検査の結びつきを重要とするため、ネットワークの項目は「患者→問題点（心身機能・活動・参加）→検査（データ）」の有向グラフを生成する（図 6.1）。ノードの色は患者ノードがグレー、心身機能の問題点が赤、活動の問題点がオレンジ、参加の問題点が黄色、検査（データ）が緑に設定する。

6.3 実験目的

本章の実験の目的は2つある。第1の目的は、5.6章で述べたようにどのような条件で初学者・熟達者ともに推論のぼらつきが生じやすいか検討することを目的とする。5章では、大腿骨転子部骨折術後の模擬症例は標準的なクリニカルパスよりも遅れており、痛みも強く動作レベルが低く設定したため、理学療法士によって重要とするべき問題点がさまざまであった。本章の実験では、模擬症例の術後の経過を良好で動作レベルが高い例と、5章と同等の経過・動作レベルの患者の段階をつけることで初学者・熟達者でどのような推論パターンや偏りを明らかにする。

第2の目的は、参加者間において頻度が高く、共有された「問題点（心身機能・活動・参加）と検査」の組み合わせを「共通解」と定義し、それらとの適合率・再現率を明らかにすることである。適合率と再現率を算出することで、各参加者の臨床推論が、共通して重要と認識されやすい判断とどの程度一致しているか、また重要とされる組み合わせをどの程度取りこぼさずに列挙できているかを定量的に評価できる。適合率は、参加者が列挙した判断のうち、どの程度が共通解に含まれているかを示す指標であり、列挙した問題点・検査の「妥当性」や「焦点化の程度」を反映すると解釈できる。一方、再現率は、共通解として定義された重要な組み合わせのうち、参加者がどの程度を網羅できているかを示す指標であり、臨床推論における「重要情報の取りこぼしの少なさ」や「判断の網羅性」を反映する。また、初学者群と熟達者群を比較することで、熟達者が共通して重視される判断を安定して捉えているのか、あるいは初学者が特定の条件下で重要な判断を見落としやすいのかといった、経験年数に基づく臨床推論の特徴を明らかにすることを目指す。

6.4 実験方法

経験年数1年目から3年目の初学者群（11名）と経験年数6年目から18年目の熟達者群（11名）に対して、症例A・B・Cの模擬症例をGoogleスプレッドシート提示し、GoogleFormでアンケート調査を行った。模擬症例は術後15日目（POD15）の大腿骨頸部骨折であり、症例Aと症例Bは同じ軽症の症例であるが症例Aには情報量の提示を制限した。症例Bでは5章と同様の検査量を提示したが、検査の異常度は軽度にし、比較的手術後の経過は良好である設定にした。杖歩行で院内を歩いており、退院までに独歩で歩けるようになることを目標と設定した。症例Aでは、症例Bで提示した情報を制限した。例えば、歩行のパフォーマンス評価である10m歩行やTime Up Go test（以下：TUG）、6分間歩行などの検査結果は「カットオフ値を達している」と具体的な数的データを取り扱わず、また医学的情報の血液検査なども「問題なし」と表示した。症例Cでは、症例A・Bに比べて手術後の経過は遅く、手術後の痛みが強く人的な見守り下で歩行器歩行で歩ける設定した（図6.2）。

GoogleFormでの質問項目は、起こりうる心身機能の問題点（最大5個まで）と1つの心身機能の問題点を導くために必要な検査（最大6個まで）と、活動の問題点（最大5個まで）と1つの活動の問題点を導くために必要な検査（最大6個まで）と、参加の問題点（最大5個まで）と1つの活動の問題点を導くために必要な検査（最大6個まで）とした。

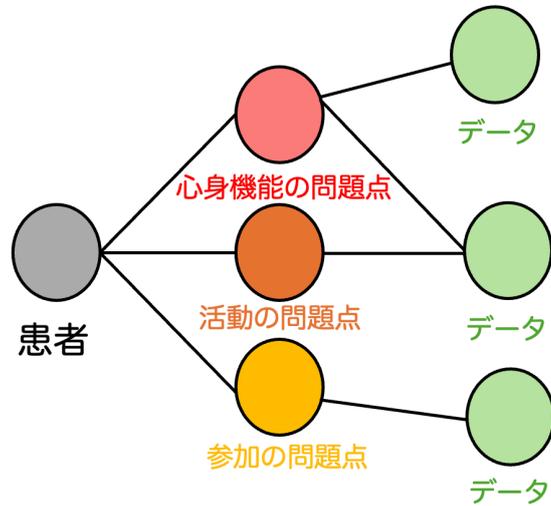


図 6.1: ネットワークの構造

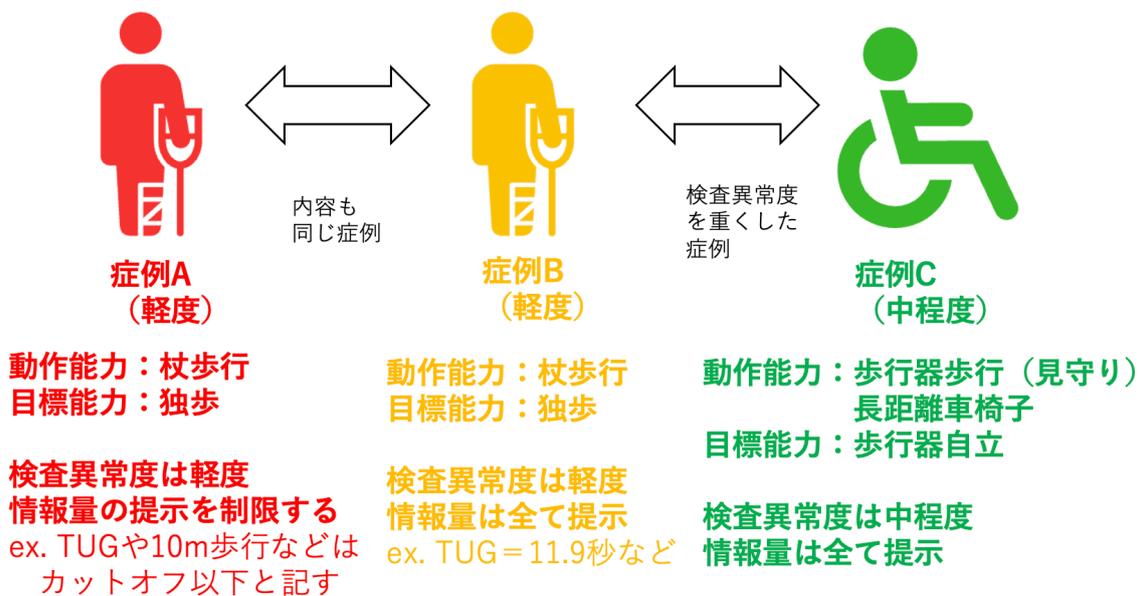


図 6.2: 模擬症例（症例 A・症例 B・症例 C）

6.5 検証と結果

6.5.1 検証1：ICF（心身機能・活動・参加）における問題点列挙の比較

症例A・B・Cでそれぞれ列挙した、心身機能・活動・参加の列挙数を初学者群・熟達者群で比較する。まずは、棒グラフで比較を行い、各群の傾向を確かめる。

まずは、症例Aの心身機能・活動・参加の結果を見る。症例Aでは、初学者群は「筋力の機能」に10人列挙しており、「関節の可動性の機能」に4人、「歩行パターン」1人という結果であった。熟達者群では、「筋力の機能」「関節の可動性の機能」「歩行パターン」のみならず、「運動耐容能」「筋の持久性機能」などと広く列挙していた（図6.3）。また症例Aの活動では、初学者群・熟達者群ともに「短距離歩行」「長距離歩行」を列挙している参加者が多く、自宅退院に向けて「屋外の歩行」「障害物を避けての歩行」など、参加者によって着目する動作が異なる傾向がみられた（図6.4）。また参加では、各群において「買い物」や「レクリエーションとレジャー」を列挙している参加者が多く、比較的ばらつきの少なかった（図6.5）。症例Bでは、心身機能・活動は症例Aと特徴的に違う点は見られなかったが、参加のみ熟達者群の方が「レクリエーションとレジャー」「買い物」「調理」を多く答えている傾向が見られた（図6.6, 図6.7, 図6.8）。

症例Cでの心身機能では、各群ともに「痛みの感覚」「筋力の機能」「関節の可動性の機能」をあげている参加者が多かったが、熟達者群では「運動耐容能」等の持久力のみならず、血液の機能をあげている参加者も見られた（図6.9）。活動では、各群多くの参加者が「短距離歩行」を列挙していたが、歩行器で歩く前段階の立つや座るなどの基本的な動作に着目する参加者もいれば、歩行器で歩く距離をより伸ばすために「長距離歩行」をあげている参加者もいた（図6.10）。参加では、熟達者群の方が「レクリエーションとレジャー」「排尿の管理」を列挙している参加者が多かったが、初学者は「衣服を着ること」に着目している参加者が多かった（図6.11）。一方、初学者群では「衣服を着ること」に着目している参加者が多かった。

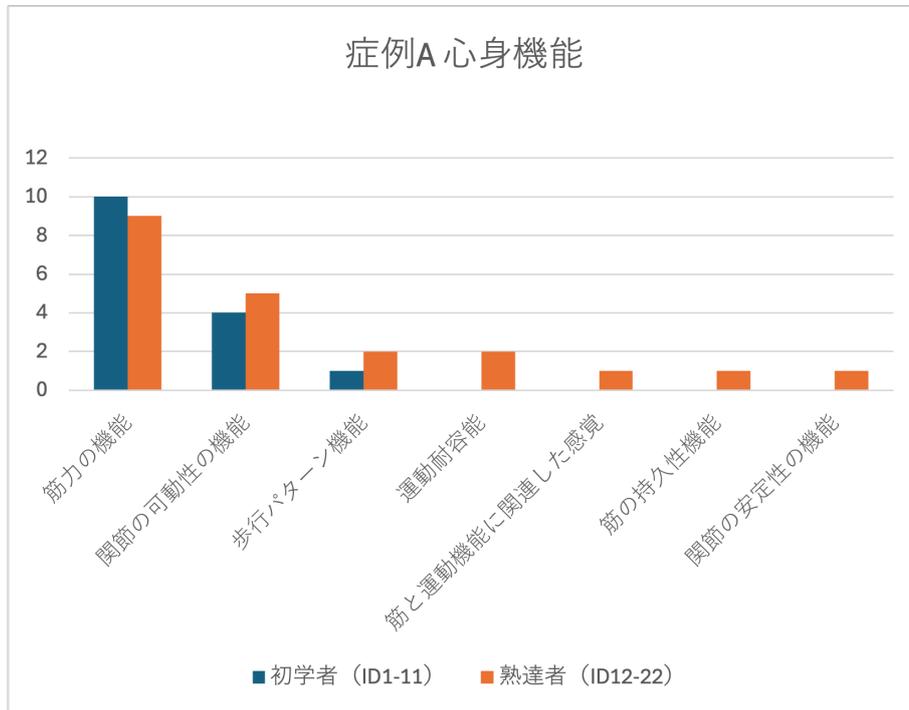


図 6.3: 症例 A の心身機能

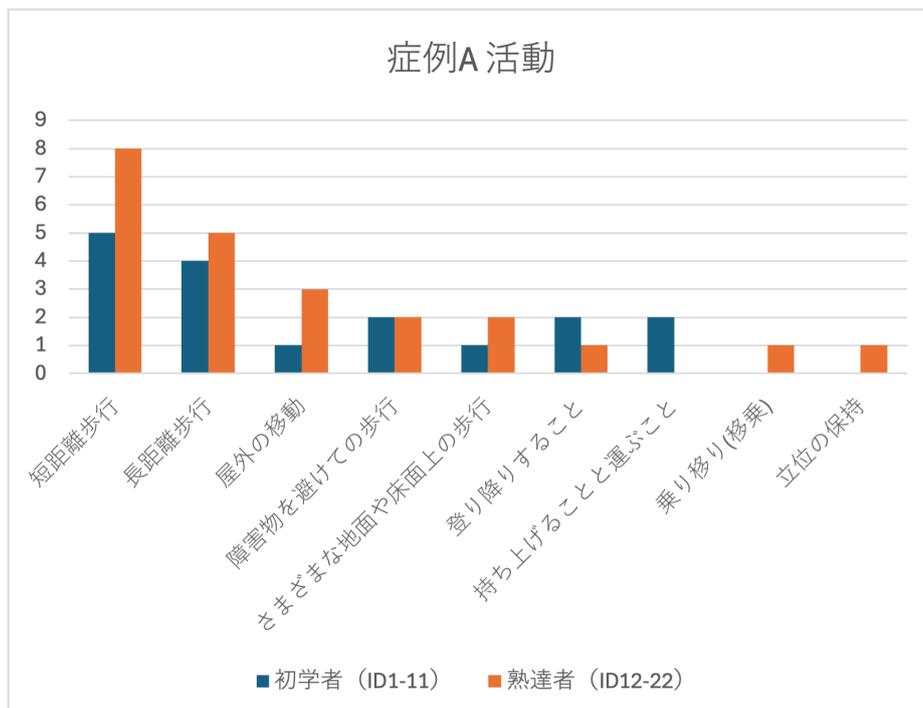


図 6.4: 症例 A の活動

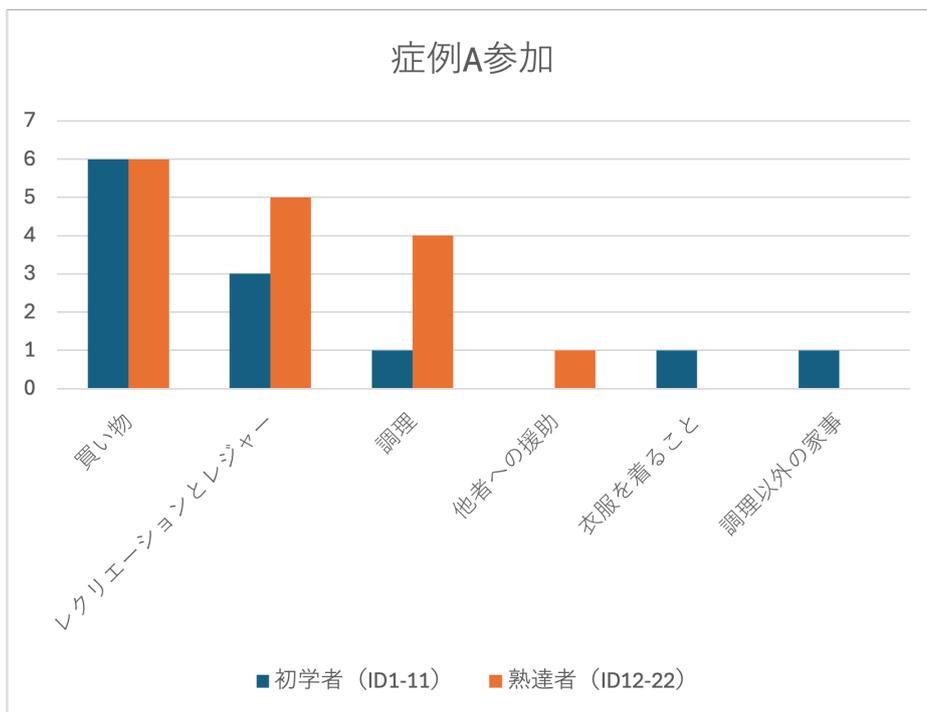


図 6.5: 症例 A の参加

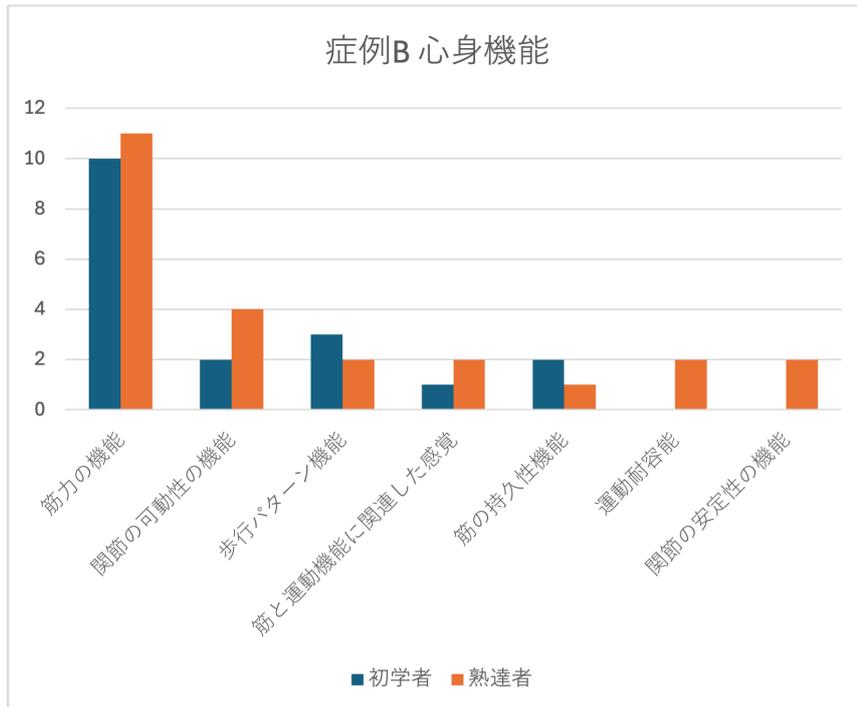


図 6.6: 症例 B の心身機能

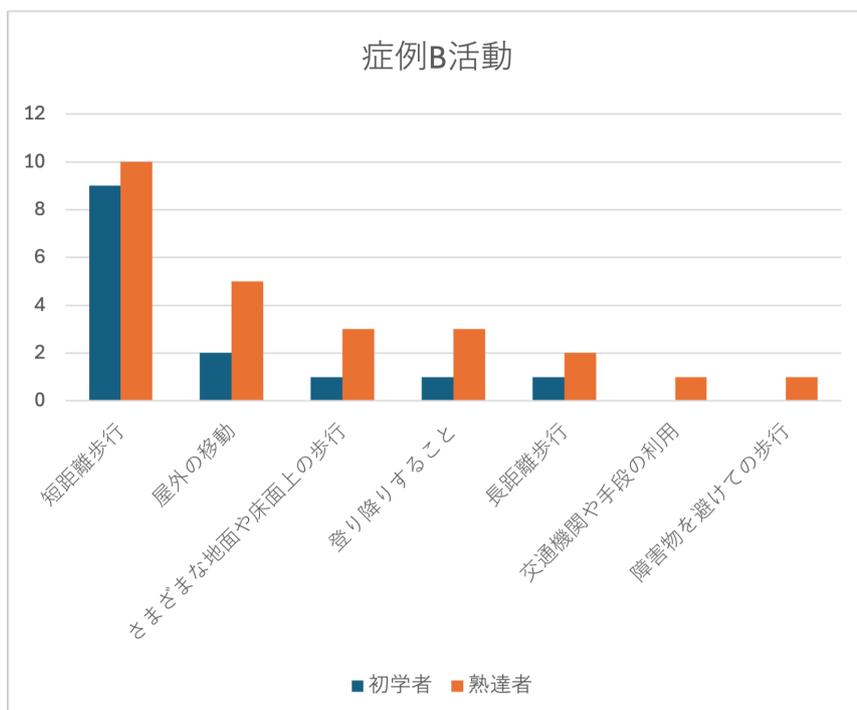


図 6.7: 症例 B の活動

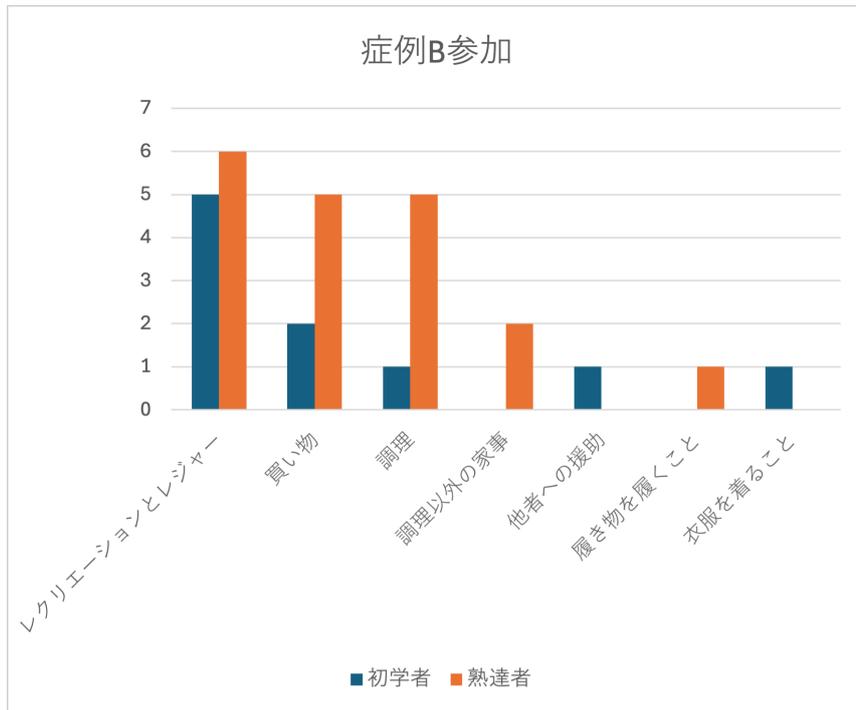


図 6.8: 症例 B の参加

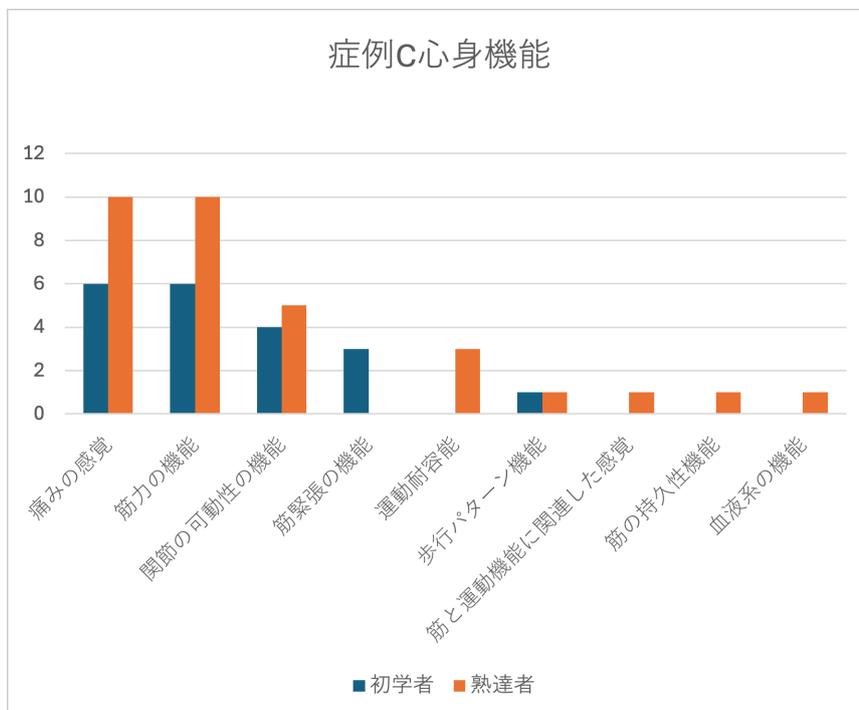


図 6.9: 症例 C の心身機能

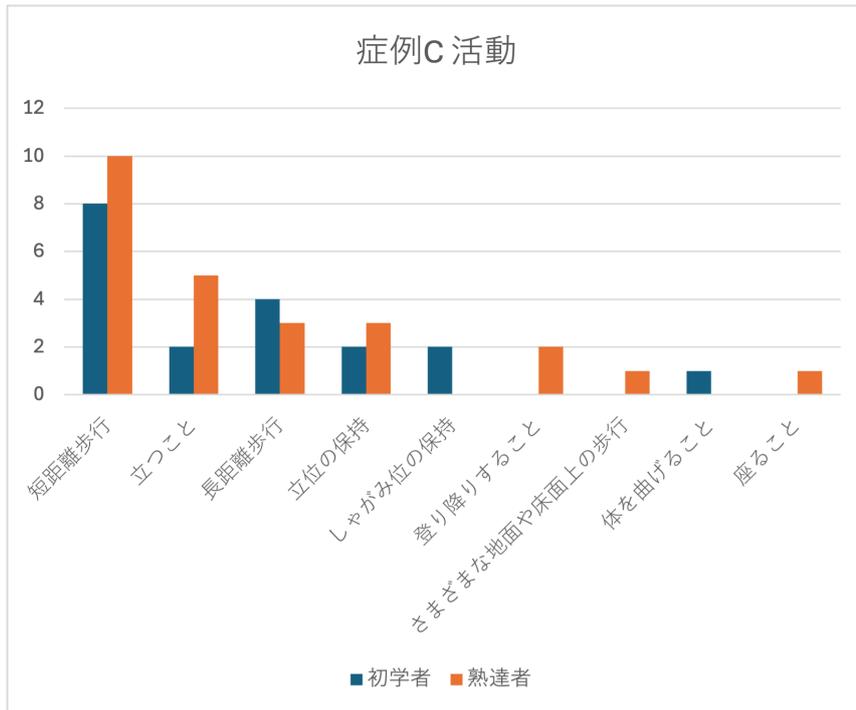


図 6.10: 症例 C の活動

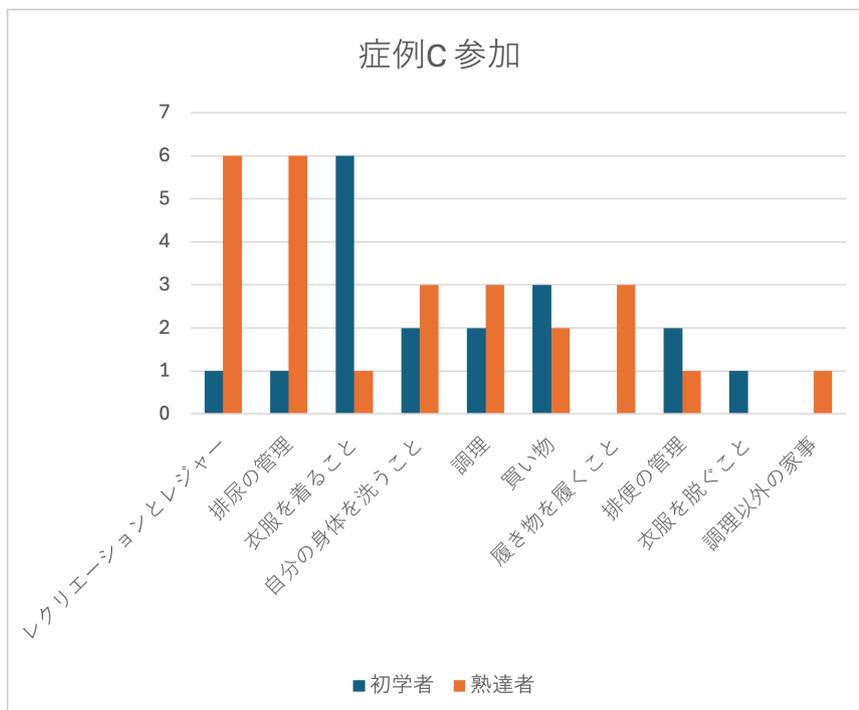


図 6.11: 症例 C の参加

6.5.2 検証 2：共通解との適合率・再現率

共通解は、症例 A・B・C に対して参加者が回答した「問題点（心身機能・活動・参加）と検査」の組み合わせのうち、頻度が高い上位 5 位とする。その結果を表 6.2, 表 6.3, 表 6.4 に示す。これらを共通解とし、各参加者の心身機能・活動・参加における適合率, 再現率を算出する。式 (6.1) に適合率を, 式 (6.2) に再現率を示す。また, 共通解の算出に当該参加者の回答が含まれないよう, Leave-one-out を適用した。

その後, 初学者群と熟達者群で適合率・再現率の群間比較を行う。結果を表 6.1 に示す。症例 C の心身機能の再現率のみ熟達者の方が有意に高く, その他は有意差を認めなかった。

$$\text{適合率} = \frac{(\text{参加者が答えた問題点・検査の組み合わせ}) \cap (\text{Top 5 共通解})}{(\text{参加者が答えた問題点・検査の組み合わせ})} \quad (6.1)$$

$$\text{再現率} = \frac{(\text{参加者が答えた問題点・検査の組み合わせ}) \cap (\text{Top 5 共通解})}{(\text{Top 5 共通解})} \quad (6.2)$$

適合率						
症例	領域	初学者 (中央値)	熟達者 (中央値)	<i>U</i>	<i>p</i>	(<i>r</i>)
A	心身機能	0.500	0.455	68.5	0.612	(0.132)
	活動	0.000	0.250	47.0	0.376	(-0.223)
	参加	0.333	0.200	72.5	0.439	(0.198)
B	心身機能	0.333	0.400	70.5	0.530	(0.165)
	活動	0.400	0.333	69.5	0.574	(0.149)
	参加	0.500	0.200	74.5	0.372	(0.231)
C	心身機能	0.375	0.273	65.0	0.796	(0.074)
	活動	0.286	0.250	69.0	0.598	(0.140)
	参加	0.333	0.200	74.0	0.391	(0.223)

再現率						
症例	領域	初学者 (中央値)	熟達者 (中央値)	<i>U</i>	<i>p</i>	(<i>r</i>)
A	心身機能	0.800	0.600	53.5	0.660	(-0.116)
	活動	0.000	0.429	50.5	0.515	(-0.165)
	参加	0.200	0.200	56.0	0.785	(-0.074)
B	心身機能	0.400	0.600	32.5	0.062	(-0.463)
	活動	0.400	0.600	40.0	0.178	(-0.339)
	参加	0.333	0.167	53.5	0.660	(-0.116)
C	心身機能	0.333	0.667	29.0	0.037	(-0.521)
	活動	0.400	0.400	56.0	0.787	(-0.074)
	参加	0.167	0.333	45.5	0.314	(-0.248)

注：太字は $p < 0.05$ を示す。

表 6.1: 症例別の適合率および再現率の Mann-Whitney U 検定

表 6.2: 症例 A における「問題点」と「検査」の組み合わせ上位（列挙率順）

順位	問題点	検査	挙げた ID 数	列挙率
心身機能				
1 位	筋力の機能	HHD（股関節外転）	17	0.773
2 位	筋力の機能	MMT 股関節外転	17	0.773
3 位	筋力の機能	歩行観察（杖歩行）	15	0.682
4 位	筋力の機能	片脚立位	10	0.455
5 位	関節の可動性の機能	ROM-t 股関節伸展	8	0.364
活動				
1 位	短距離歩行	MMT 股関節外転	8	0.364
1 位	短距離歩行	TUG（独歩）	8	0.364
2 位	短距離歩行	歩行観察（杖歩行）	7	0.318
3 位	短距離歩行	10m 歩行（独歩）	6	0.273
3 位	短距離歩行	主訴	6	0.273
3 位	短距離歩行	希望獲得 ADL	6	0.273
3 位	短距離歩行	独歩	6	0.273
参加				
1 位	買い物	自宅周りの環境	8	0.364
2 位	買い物	屋外歩行	7	0.318
2 位	買い物	趣味	7	0.318
3 位	レクリエーションとレジャー	趣味	6	0.273
3 位	買い物	元々の移動能力	6	0.273

表 6.3: 症例 B における「問題点」と「検査」の組み合わせ上位（列挙率順）

順位	問題点	検査	挙げた ID 数	列挙率
心身機能				
1 位	筋力の機能	MMT 股関節外転	19	0.864
2 位	筋力の機能	歩行観察（杖歩行）	14	0.636
3 位	筋力の機能	HHD（股関節外転）	13	0.591
4 位	筋力の機能	主訴	7	0.318
5 位	関節の可動性の機能	ROM-t 股関節伸展	6	0.273
活動				
1 位	短距離歩行	独歩	14	0.636
2 位	短距離歩行	MMT 股関節外転	11	0.500
2 位	短距離歩行	歩行観察（杖歩行）	11	0.500
4 位	短距離歩行	TUG（独歩）	10	0.455
5 位	短距離歩行	Berg Balance Scale	8	0.364
参加				
1 位	レクリエーションとレジャー	趣味	9	0.409
2 位	レクリエーションとレジャー	屋外歩行	7	0.318
3 位	買い物	屋外歩行	6	0.273
4 位	レクリエーションとレジャー	自宅周りの環境	5	0.227
4 位	調理	家事	5	0.227
4 位	買い物	元々の移動能力	5	0.227

表 6.4: 症例 C における「問題点」と「検査」の組み合わせ上位（列挙率順）

順位	問題点	検査	挙げた ID 数	列挙率
心身機能				
1 位	痛みの感覚	NRS（荷重時）	14	0.636
2 位	筋力の機能	MMT 股関節外転	13	0.591
3 位	痛みの感覚	NRS（ROM 時）	11	0.500
4 位	痛みの感覚	炎症所見	10	0.455
5 位	痛みの感覚	MMT 股関節外転	9	0.409
5 位	痛みの感覚	主訴	9	0.409
活動				
1 位	短距離歩行	歩行器歩行	12	0.545
2 位	短距離歩行	BI 歩行	10	0.455
2 位	短距離歩行	NRS（荷重時）	10	0.455
4 位	短距離歩行	歩行観察（杖歩行）	9	0.409
5 位	立つこと	立ち上がり	6	0.273
参加				
1 位	レクリエーションとレジャー	趣味	6	0.273
1 位	衣服を着ること	BI トイレ	6	0.273
3 位	排尿の管理	BI トイレ	5	0.227
3 位	排尿の管理	立位保持	5	0.227
3 位	自分の身体を洗うこと	BI 入浴	5	0.227
3 位	衣服を着ること	BI 着替え	5	0.227

6.5.3 検証3：集約ネットワークの比較

検証3では、症例A・B・Cの比較を熟達者群・初学者群の集約ネットワークの比較を行う。前章の検証4と同じように比較を行う。ノード・エッジ表現も同じように、「問題点」「検査」の2つの組み合わせを参加者が5人以上挙げていると赤エッジ、参加者が4人挙げていると青エッジ、参加者が3人挙げている場合は緑エッジとした。また参加者が2人以下の組み合わせは非表示とした。また赤エッジは、5人以上で表示されるため、エッジ上に回答した人数を表示した。

症例Aの結果を図6.12と図6.13に示す。まずは、熟達者群に着目する(図6.12)。赤エッジは、心身機能の「筋力の機能」から「歩行観察」「MMT 股関節外転」「HHD 股関節外転」「片脚立位」へ、「関節の機能」からは「ROM 股関節伸展」へ伸びている。また、活動の「短距離歩行」から歩行パフォーマンス評価の「TUG (杖)」へ伸びており、参加では、「買い物」から「自宅周りの環境」へエッジが伸びていた。初学者群(図6.13)では、赤エッジは心身機能の「筋力の機能」から「歩行観察」「MMT 股関節外転」「HHD 股関節外転」「ROM 股関節内転」「片脚立位」へエッジが伸びている。熟達者群・初学者群で比較すると、まず多くの参加者が着目している問題点のノード数が少ない。また初学者群は赤エッジも心身機能の「筋力の機能」からしか出ておらず、多くの初学者が筋力が問題だと答えている傾向がある。熟達者群では、心身機能・活動・参加の3つの領域ともに検査に向かって赤エッジが出ており、心身機能のみならず幅広く問題点や検査を列挙していた。次に症例Bの結果を図6.14と図6.15に示す。模擬症例の設定としては、症例Aと同じであるため、ネットワーク自体に症例Aのネットワークと似た傾向であった。

症例Cでは、症例A・Bとは異なり、検査の異常度を高めたため、初学者群・熟達者群ともにネットワーク自体が大きくなった。まずは熟達者群に着目する(図6.14)。熟達者群では、「痛みの感覚」から「NRS (荷重時痛)」「MMT 股関節外転」「主訴」「炎症」「NRS (ROM時)」「大腿周径」と痛みや炎症所見、主訴との結びつけを挙げる人が多かった。他にも「筋力の機能」から「NRS (荷重時痛)」「MMT 股関節外転」「HHD 股関節外転」と股関節外転に着目している参加者が多かった。また活動では「短距離歩行」「立つこと」「立位の保持」の基本的な動きを挙げている参加者もいれば、「長距離歩行」などの応用動作を上げている参加者もいた。参加も「排尿の管理」「自分の体を洗うこと」「履き物を履くこと」などのセルフケアに着目する参加者もいれば、「レクリエーションとレジャー」のように退院後の趣味活動に着目しているものもいた。初学者群も同じように赤エッジに着目すると(図6.15)、心身機能では「痛みの感覚」から「NRS (荷重時痛・ROM時)」や「炎症」などの痛みや炎症に着目しており、「筋力の機能」と「MMT 股関節外転」が結びついており、熟達者群と傾向は似ていた。しかし、活動は「短距離歩行」や「長距離歩行」と歩行に着目した参加者が多かった。また、参加ではセルフケアにあたる「衣服を着ること」を答えている参加者が多く、熟達者を比べると退院後の問題点よりも現状の課題を優先的に考える参加者が多かった。

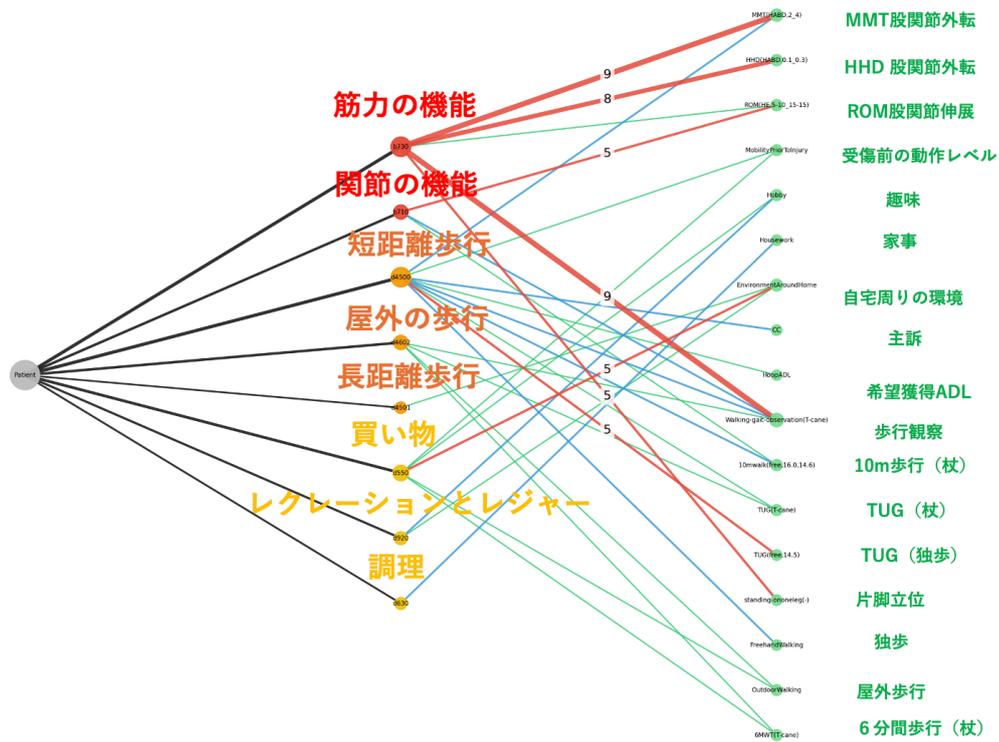


図 6.12: 症例 A の熟達者の集約ネットワーク

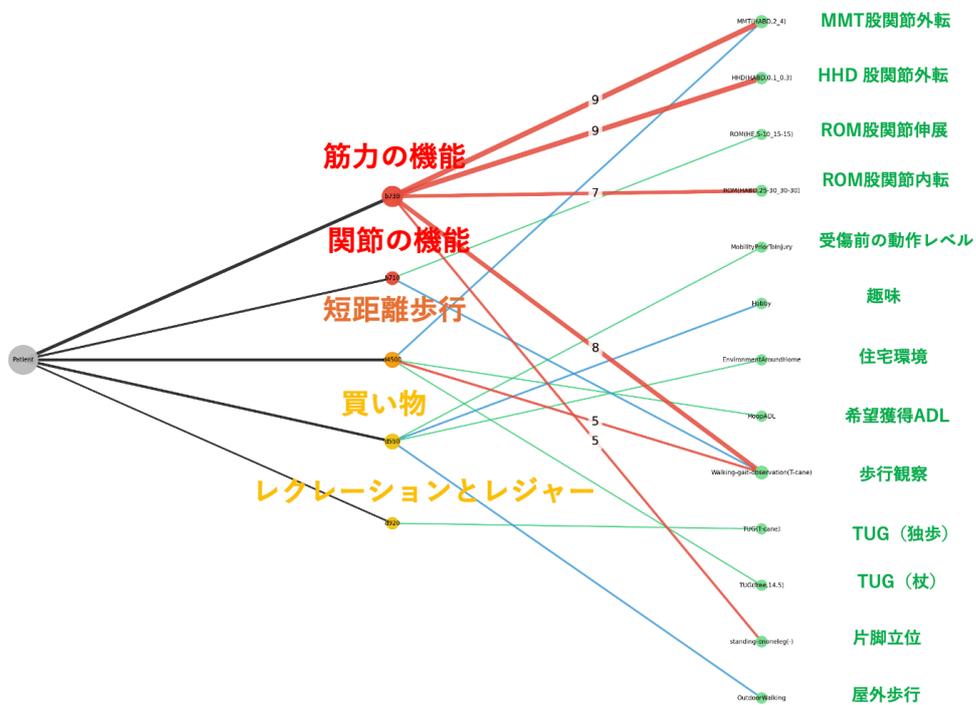


図 6.13: 症例 A の初学者の集約ネットワーク

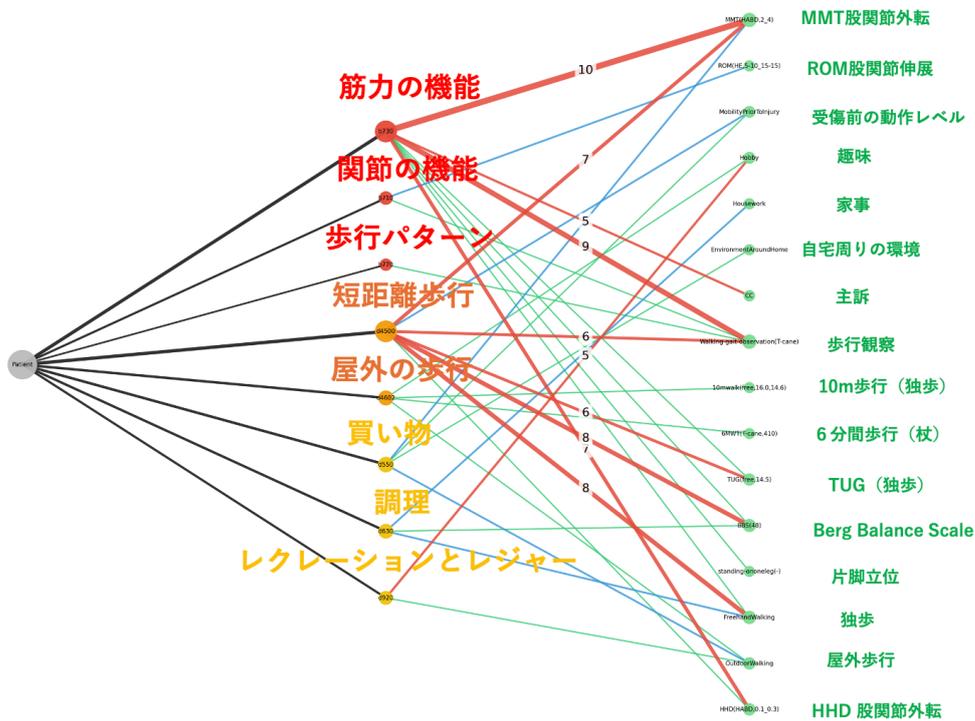


図 6.14: 症例 B の熟達者の集約ネットワーク

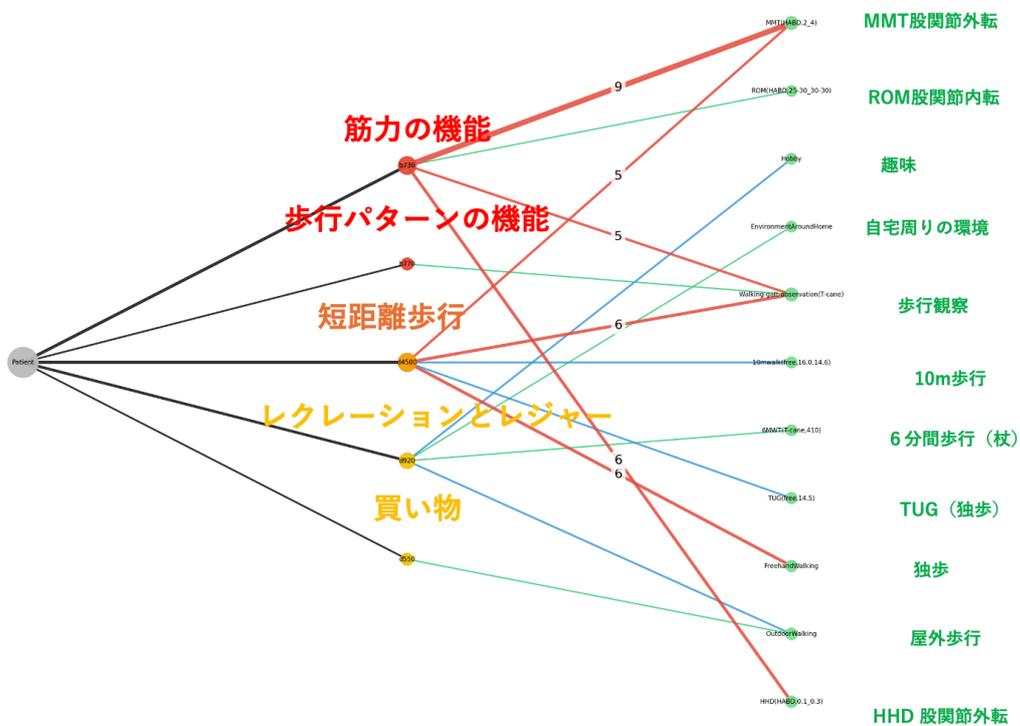


図 6.15: 症例 B の初学者の集約ネットワーク

6.6 本章の考察

検証1・2・3の結果に基づき考察する。6.5.1章の検証1では、ICF（心身機能・活動・参加）における問題点の列挙について、初学者群と熟達者群の比較を行った。その結果、症例A・B・Cのいずれにおいても、熟達者群の方が列挙数が多く、より幅広い問題点を挙げる傾向が示された。特徴的であったのは、症例Cの「参加」における問題点の列挙である。症例Cは、術後経過が標準的なクリニカルパスより遅延しており、疼痛が強く、歩行器見守りで歩行可能な状態として設定した。この条件において、初学者群では「衣服を着ること」を問題点として挙げる者が多かった一方、熟達者群では「レクリエーションとレジャー」および「排尿の管理」の2つに回答が分かれた。この差異は、目標設定の時間軸の違いを反映している可能性がある。すなわち、初学者群は短期的な到達目標として「参加」の問題点を列挙する傾向がみられたのに対し、熟達者群では、現時点で実施困難な活動を短期的な課題として捉える参加者と、趣味活動への参加といった長期的な目標を見据えて問題点を挙げる参加者が存在したと考えられる。

6.5.2章の検証2では、症例A・B・Cに対する回答から、「問題点（心身機能・活動・参加）と検査」の組み合わせのうち列挙頻度が高い上位5位を「共通解」と定義し、各参加者の回答との適合率および再現率を算出した。ここで用いた共通解は、専門家の合意に基づく正解ではなく、参加者集団において合意が高い組み合わせの集合である。したがって本検証では、初学者群・熟達者群の臨床推論が、参加者集団で重要と認識されやすい判断とどの程度一致しているか、ならびに重要とされる組み合わせをどの程度取りこぼさずに列挙できているかを算出した。

検証2の結果、症例Cの心身機能における再現率のみ、熟達者群が有意に高かった一方、症例A・Bでは群間に有意差は認められなかった。症例A・Bでは、術後経過を良好に設定した影響により、初学者群と熟達者群の回答傾向が大きく分岐しなかった可能性がある。検証1においても、症例A・Bでは熟達者群の方が列挙数は多いものの、問題点の種類や着目領域の傾向は初学者群と大きくは異ならなかった。これらを踏まえると、少なくとも提示情報量が臨床推論に与える影響は経験年数によって大きく変化しなかったと考えられる。一方、症例Cでは、5章のPOD7と同程度の動作レベルに相当するように設定したところ、心身機能における再現率のみ熟達者群が高かった。このことは、症例の検査異常度が高い条件では、熟達者群の方が参加者集団で重要と共有されやすい組み合わせの取りこぼしが少なかった可能性を示唆する。また症例Cは、模擬症例の到達目標能力を歩行器自立レベルとし、症例A・Bと比較して疼痛が強く動作レベルも低い条件であったため、問題点および検査の列挙数が全体として減少した（表6.4）。集約ネットワークを確認すると、症例Cの熟達者群では、心身機能の共通解上位5位の全てに対応するエッジが多く観察された。以上より、熟達者群は共有されやすい重要な組み合わせを網羅的に回答する傾向があり、取りこぼしが少なかったことが再現率の上昇に寄与したと考えられる。

加えて、認知心理学面的側面から考察を述べる。Normanらは、診断エラーの原因として直感の心理（システム1）に由来する認知バイアスが強調されてきた一方で、実証研究においては、知識不足や知識構造の未熟さが依然として重要な要因であることを示している [10]。すなわち、診断エラーの対策は単に分析的に熟慮すること（システム2）やバイアスを意識

することによって向上するのではなく、知識量や知識を構造する思考能力に大きく依存することを示唆する。Schmidtらの医療における専門性の発達研究では、初学者は詳細な因果関係に基づいて、順序立てて推論を行う傾向がある一方、熟達者は知識がカプセル化され、「疾病スクリプト」として再構造化された知識を活用することで、迅速かつ効率的に仮説を行うとされている[11]。本研究において熟達者群が共通解の取りこぼしを減少させた結果は、「疾患スクリプト」が活用され、臨床推論の網羅性を高めた可能性がある。さらに、熟達者の迅速な意思決定を説明する理論として、Kleinの「Recognition-Primed Decisionモデル（以下：RPDモデル）」が挙げられる[8]。RPDモデルでは、熟達者は複数の選択肢を比較し、検討するのではなく、現在の状況を過去の経験と照らし合わせる。そして、典型的なパターンとして認識することで、最初に想起された行動を内的にシミュレーションし、実行可能であれば即座に採用するとされる[8]。すなわち、熟達者の判断は単なる直感ではなく、経験に基づいて構造化された知識に支えられた認知過程である。RPDモデルの観点からは、熟達者は状況を意味あるパターンとして把握し、重要とされる情報を選択的に抽出することで、判断の網羅性と効率性を両立させていた可能性がある。以上より、本研究の結果は、診断エラーを単なる認知バイアスの問題として捉えるのではなく、知識構造の蓄積、疾病スクリプトの形成、および専門性の発達過程に基づくと整合的であると考えられる。

6.7 本研究の展望

4章の結果より、アセスメント文から臨床推論を構造化し、ネットワークとして表現することで、理学療法士間の推論を客観的に比較できる可能性が示された。また5章の結果より、集約ネットワークによる比較では、初学者群は熟達者群に比べてPOD7からPOD30にかけても着目点が大きく変化せず、着目範囲が限定的であることが示唆された。さらに6章の結果から、熟達者群は参加者集団で重要と共有されやすい問題点・検査の組み合わせについて取りこぼしが少なく、より網羅的に回答する傾向が得られた。

以上の知見は、初学者理学療法士に対して、問題点および検査の見逃しを低減する観点から、理学療法プロセスにおける「②情報収集」および「③評価」の段階での支援が重要であることを示唆する。一方で、本研究の結果は、単に「列挙数を増やす」ことを目的とするものではない。臨床推論において重要となるのは、症例の病期やリスクに応じて、必要な情報を過不足なく収集し、それらを問題点へ適切に結び付けることである。したがって今後は、初学者が「何を見ればよいか」だけでなく、「なぜそれを見るのか」「その情報をどの問題点にどう統合し、解釈するのか」という推論過程を明示化し、教育やリフレクションに接続する支援が求められる。

この課題に関連し、岩脇らは理学療法初学者の情報収集支援を目的とした患者情報記録ツールを提案している[14]。患者情報記録ツールは、情報収集や理学療法評価の記録を支援する一方で、どの情報や評価がどの問題点と結び付けられたのかといった「④統合と解釈」および「⑤問題点の抽出」の思考過程は明示されにくい。そこで本研究の展望として、本研究で用いたネットワーク表現を、理学療法士が自身の思考過程をリフレクションできるツールやシステムへ応用することが挙げられる。具体的には、理学療法士が選択した検査、得られた所見、抽出した問題点、およびそれらを結び付けた根拠を、時間経過（POD7、

POD30 など) とともに同一の枠組みで可視化し、推論の変化や偏りを自覚できるようにすることが考えられる。

初学者のこのような支援のプロダクト開発を行い、行動変容デザインを考える必要がある。行動デザインには、4つの段階がある [12]。1つ目が、心の動きが行動の決定にどう作用するのか、またそれが行動の変化にどう関係するのかを「理解する」。2つ目が企業やユーザーの目標を踏まえて、変えるべき行動を「探索する」。3つ目が設定した行動に向けて、プロダクトを「デザインする」。4つ目が慎重な測定と分析に基づいて、プロダクトの効果を反復的に「改善する」。本研究では、1つ目の「理解する」、2つ目の変えるべき行動を「探索する」ことができた。今後は、初学者が見落としを減らせるように「デザイン」する必要がある。臨床推論をネットワークとして可視化することで、例えば実施した検査の偏りや、問題点への結び付けの不足といった推論過程の特徴を自覚できる可能性がある。今後は、この可視化を教育・フィードバックに接続し、初学者の情報収集および評価の質向上に資する支援方法の検討が必要である。

7 結論

本研究の主たる課題は、理学療法士の臨床推論が暗黙知として扱われやすく、その思考過程の再現性および説明可能性が十分に担保されていないために、理学療法士間で客観的に比較することが困難である点にある。この課題は、養成校および臨床現場における臨床推論教育の体系化を妨げ、初学者理学療法士が臨床推論に苦手意識を抱く一因となっている。そこで本研究では、臨床推論の思考過程を比較可能な形式へ変換し、臨床推論教育に資する知見を得るとともに、初学者理学療法士に必要な支援の方向性を明らかにすることを目的とした。以上を踏まえ、「臨床推論構造をネットワークで可視化することにより、経験年数および患者情報の差異に応じて推論構造はどのように異なるのか」というRQを提起した。本研究では、上記RQを段階的な検証課題に分解し、各章の実験により検証した。

4章では、アセスメント文を構造化し、ネットワークで可視化することで、理学療法士間の推論を客観的に比較することは可能かを検証した。その結果、アセスメント文同士で比較するよりも、客観的に比較可能となった。また構造化によって理学療法の対象となる問題点を抽出することができ、形式的かつ定量的に評価が可能となった。定量的評価の結果から、臨床経験によって問題点を絞り込む能力は向上するが、一方で馴化に伴い問題点を見落とす傾向も生じ得るため、経験年数が増加しても臨床推論の的確さが一様に向上するとは限らない可能性が示唆された。

5章では、急性期・回復期の2時点の患者情報を提示することにより、初学者・熟達者の臨床推論ネットワークにどのような差異がみられるのかを検証した。その結果、初学者群では急性期から回復期にかけてノード数が有意に減少し、術後経過に伴いネットワークが収束しやすい傾向が示された。しかし、初学者群の集約ネットワークでは、多くの参加者が急性期から回復期にかけて「筋力」に主として着目しており、患者情報・検査結果を変化させたにも関わらず着目範囲は限定的であった。一方で熟達者群では、ノード数・エッジ数に有意な差は認められなかったが、集約ネットワークでは「筋力」に加えて「痛み」「関節可動域」など複数の観点から検査や問題点に着目できていた。なお、5章の所見は4章と一部整合しなかったが、これはアセスメント文と選択式回答というデータ取得方法の違いにより、顕在化する推論過程の表出が異なった可能性があると考えられる。また5章のリミテーションとして、模擬症例の重要な問題点・検査を定めなかったため、参加者の推論の妥当性を直接判断できなかった。そのため6章では、専門家同意による正解ではなく、参加者で高頻度を選択された上位5位の組み合わせを「共通解」として定義し、適合率・再現率による評価を行った。

6章では、提示情報量および検査異常度の差異が臨床推論ネットワークおよび共通解に及ぼす影響を検証した。その結果、症例Cの心身機能における再現率のみ熟達者群が有意に高く、熟達者群が実験参加者間で重要と共有されやすい組み合わせを見落としにくい可能性が示された。以上より、提示する患者情報量が少なく、かつ検査異常度が低い条件では経験年数による差は大きく現れなかった一方、検査異常度が高い条件では、熟達者群の方が重要と共有されやすい「問題点」と「検査」の組み合わせを取りこぼしにくい可能性が示唆された。

以上の結果から、本研究で提案した臨床推論の構造化・ネットワーク可視化は、思考過

程を比較可能な形式へ変換し、理学療法士間の推論の共通点と差異を客観的に捉える可能性を示した。また、病期変化や提示情報量・検査異常度の差異により、初学者群と熟達者群で着目点や推論の網羅性が異なることが示唆された。特に初学者群では着目範囲が限定的で見落としが生じやすい一方、熟達者群では重要と共有されやすい問題点・検査の組み合わせを取りこぼしにくい傾向が得られた。加えて、認知心理学的観点からは、熟達者群の取りこぼしの少なさは、認知バイアスの制御のみならず、知識量や知識構造の体系化が診断精度に関与するという見解と整合する。疾病スクリプトやRPDモデルの観点からは、熟達者は状況をパターンとして把握し、重要情報を選択的に抽出することで、網羅性と効率性を両立させていた可能性がある。

最後に、本研究の知見を踏まえ今後の課題を述べる。初学者理学療法士に対しては、理学療法プロセスにおける「②情報収集」および「③評価」の段階で、必要な情報を過不足なく収集し、問題点へ適切に結び付けるための支援が重要であると考えられる。この課題に関連し、岩脇らは理学療法初学者の情報収集支援を目的とした患者情報記録ツールを提案している[14]。一方で、情報や評価がどの問題点と結び付けられたのかといった「④統合と解釈」および「⑤問題点」の思考過程は明示されにくい。そこで今後は、本研究で用いた臨床推論の構造化とネットワーク可視化を、理学療法士が自身の推論の偏りや不足をリフレクションできるツールやシステムへ応用し、教育・フィードバックに接続することを目指す。

8 謝辞

本研究の遂行および本修士論文の執筆にあたり、関西大学大学院総合情報学研究科の松下光範教授には様々なご指導ご鞭撻を賜りました。私は修士課程からの入学だったにも関わらず、多くの経験をさせていただきました。入学して1ヶ月も経たない間に、世界理学療法アジア西太平洋地域学会の抄録提出に関してご指導いただき、2024年9月にインドネシア・バリ島で初の学会発表を達成することができました。初の学会発表だったにも関わらず、多くの質問を受け、日本のみならずアジア各国の理学療法士と交流ができました。そこから私の人生が少しずつ自信がつき始め、研究者としての道が始まりました。学会活動のみならず、松下教授には学問に対する楽しさを教わりました。理学療法のみならず、器・コミック・ニュース分野などを学び、私の日常生活で触れるものが豊かになりました。修士論文・学会の抄録・予稿などでは大変ご迷惑をおかけしましたが、修士課程の2年間は人生にとってかけがえのないものになりました。今後も、情報学×理学療法の先駆者の一員として駆け抜けていけるように精進していきます。心より深く感謝を申し上げます。

また甲南女子大学看護リハビリテーション学部理学療法学科の堀寛史教授にも多くのご指導ご鞭撻を賜りました。堀教授には藍野大学の3年生頃から授業以外でも、大学院進学についてご教授いただきました。社会人2年目の時に松下教授をご紹介いただき、社会人3年目で本研究室に入学できました。修士課程では病院の常勤勤務との両立で不安を抱き、自信をなくすことが多く、自分は理学療法も研究も向いていないんじゃないかと何度もご相談させていただきました。その際に漫画「3月のライオン」のワンシーンを送っていただき、「しつこくてあきらめられない気持ちを『向いてる』というんだよ」と言葉をかけてくださりました。何度も諦めかけた道で背中を押し続けてくださったのは堀教授でした。今後は、私も多くの人を助けられる理学療法士になれるように精進いたします。心より深く感謝を申し上げます。

藍野大学医療保健医療学部理学療法学科の杉本明文先生、本研究室の卒業生である、神戸市立医療センター中央市民病院の畠山駿弥先生、岸和田徳洲会病院の吉田龍洋先生にも大変お世話になりました。理学療法チームの定例会議では、至らなかった場面もあり、多くのご指導をいただきました。杉本先生とは国際学会や国内の理学療法学会で多くのサポートをいただきました。また畠山先生、吉田先生には実験協力等していただき、本修士論文が完成いたしました。心より深く感謝を申しあげます。今後も研究だけでなく、理学療法士としてもご指導賜りますと幸いですよろしく願いいたします。

また本研究室の唯一の同期である高橋りさ氏には大変お世話になりました。いつも冷静で優しく、寄り添ってくれたおかげで、乗り越えれたものが多くありました。また卒業してもたくさん遊んでください。また本研究室の理学療法チームの三藤優実氏、東岡秀樹氏にはお世話になりました。研究進捗報告では多くの刺激をもらい、研究を楽しくかつ前向きに進めることができましたと思います。最後に仕事と学生生活で多忙な時も常に支えてくださった母に感謝の意を表すとともに謝辞といたします。

研究業績

- Kanae Takahashi, Shunya Hatakeyama, Tatsuhiro Yoshida, Mitsunori Matsushita, Hirofumi Hori. Relationship between physical therapy examination skills and clinical reasoning of novice physiotherapists, World Physiotherapy Asia Pacific Region Congress 2024, No.2-2, Poster, 2024.
- 高橋 可奈恵, 畠山 駿弥, 吉田 龍洋, 堀 寛史, 松下 光範. 初学者理学療法士の臨床推論教育を目指したアセスメント文の構造化に関する検討, 電子情報通信学会 HCG シンポジウム 2024 論文集, No.C-6-2, 2024.
- 高橋 可奈恵, 畠山 駿弥, 吉田 龍洋, 堀 寛史, 松下 光範. 部分支援による情報収集が臨床推論力と問題点抽出力に及ぼす効果, 第 13 回日本理学療法教育学会学術大会, No.P2-1-4, 2024.
- 高橋 可奈恵, 畠山 駿弥, 吉田 龍洋, 堀 寛史, 松下 光範. 統合と解釈文の構造的可視化による理学療法推論能力の直観的把握の試み, 第 1 回日本理学療法推論学会学術集会, 2025.
- 高橋 可奈恵, 三藤 優実, 堀 寛史, 松下 光範. 理学療法推論時の検査適合率と経験年数の関係—テキストデータをデジタル処理するための基礎研究—, 第 3 回デジタル理学療法研究会学術大会, 2025.
- 岩脇 朱梨, 高橋 可奈恵, 堀 寛史, 松下 光範. 理学療法初学者の臨床情報収集支援を目的とした患者情報記録ツールの提案, ヒューマンインタフェース学会論文誌, Vol.27, No.3, pp.165-176, 2025.
- Kanae Takahashi, Hirofumi Hori, Mitsunori Matsushita. A Network Analysis of Clinical Reasoning after Femoral Transverse Fracture Surgery: A Comparison by Years of Experience, *Formosan Journal of Physical Therapy*, Vol.50, No.4, pp.320-321, 2025.
- 高橋 可奈恵, 堀 寛史, 松下 光範. ネットワーク分析による理学療法推論の可視化—教育支援のためのシステム開発に向けて—, 第 14 回日本支援工学理学療法学会学術大会, O-13-2, 2025.

参考文献

- [1] Abraham, S. M. and Sudhamathy, G.: Mapping Knowledge: Evaluating Concept Maps Using Social Network Analysis and Network Similarity Measures, *Journal of Theoretical and Applied Information Technology*, Vol. 103, No. 12, pp. 5374–5394 (2025).
- [2] Charlin, B., Brailovsky, C. A., Leduc, C. and Blouin, D.: The diagnostic script questionnaire: A new tool to assess a specific dimension of clinical competence, *Advances in Health Sciences Education: Theory and Practice*, Vol. 3, pp. 51–58 (1998).
- [3] Cohen, J. F., Korevaar, D. A., Altman, D. G., Bruns, D. E., Gatsonis, C. A., Hooft, L., Irwig, L., Levine, D., Reitsma, J. B., de Vet, H. C. W. and Bossuyt, P. M. M.: STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration, *BMJ Open*, Vol. 6, No. 11, p. e012799, DOI: 10.1136/bmjopen-2016-012799 (2016).
- [4] Croskerry, P.: From Mindless to Mindful Practice — Cognitive Bias and Clinical Decision Making, *The New England Journal of Medicine*, Vol. 368, No. 26, pp. 2445–2448, DOI: 10.1056/NEJMp1303712 (2013).
- [5] Croskerry, P., Singhal, G. and Mamede, S.: Cognitive debiasing 1: origins of bias and theory of debiasing, *BMJ Quality & Safety*, Vol. 22, No. Suppl 2, pp. ii58–ii64, DOI: 10.1136/bmjqs-2012-001712 (2013).
- [6] Evans, J. S. B. T. and Stanovich, K. E.: Dual-Process Theories of Higher Cognition: Advancing the Debate, *Perspectives on Psychological Science*, Vol. 8, No. 3, pp. 223–241, DOI: 10.1177/1745691612460685 (2013).
- [7] Gottlieb, M., Caretta-Weyer, H., Chan, T. M. and Humphrey-Murto, S.: Educator’s blueprint: A primer on consensus methods in medical education research, *AEM Education and Training*, Vol. 7, p. e10891, DOI: 10.1002/aet2.10891 (2023).
- [8] Klein, G. A.: Recognition-Primed Decision (RPD) Model, Technical Report ADA309570, U.S. Army Research Institute for the Behavioral and Social Sciences (1996). Available at Defense Technical Information Center (DTIC).
- [9] Lubarsky, S., Dory, V., Duggan, P., Gagnon, R. and Charlin, B.: Script concordance testing: From theory to practice: AMEE Guide No. 75, *Medical Teacher*, Vol. 35, No. 3, pp. 184–193, DOI: 10.3109/0142159X.2013.760036 (2013).
- [10] Norman, G. R., Monteiro, S. D., Sherbino, J., Ilgen, J. S., Schmidt, H. G. and Mamede, S.: The Causes of Errors in Clinical Reasoning: Cognitive Biases, Knowledge Deficits, and Dual Process Thinking, *Academic Medicine*, Vol. 92, No. 1, pp. 23–30, DOI: 10.1097/ACM.0000000000001421 (2017).
- [11] Schmidt, H. G. and Boshuizen, H. P. A.: On Acquiring Expertise in Medicine, *Educational Psychology Review*, Vol. 5, No. 3, pp. 205–221 (1993).

- [12] Wendel, S.: *Designing for behavior change: Applying psychology and behavioral economics*, O'Reilly (2013).
- [13] 秋田直繁, 森田昌嗣, 椎塚久雄: デザイナーの推論過程のフレームワークとその可視化に関する基礎的考察 —メタ認知補助としての推論マッピング法の提案—, 日本感性工学会論文誌, Vol. 16, No. 4, pp. 375–384, DOI: 10.5057/jjske.TJSKE-D-17-00019 (2017).
- [14] 岩脇朱梨, 高橋可奈恵, 堀寛史, 松下光範: 理学療法初学者の臨床情報収集支援を目的とした患者情報記録ツールの提案, ヒューマンインタフェース学会論文誌, Vol. 27, No. 3, pp. 165–176 (2025).
- [15] 宇都雅輝, 鈴木宏昭, 植野真臣: Toulmin モデルのベイジアンネットワーク表現を用いた論証推敲支援システム, 電子情報通信学会論文誌 D, Vol. J96-D, No. 4, pp. 998–1011 (2013).
- [16] 大住崇之: 理学療法士の卒後教育についての一考察～卒後教育の現状と課題～, 理学療法教育, Vol. 1, pp. 26–37 (2022).
- [17] 大西弘高: 臨床推論の評価法, 日本内科学会雑誌, Vol. 97, No. 10, pp. 2596–2603, DOI: 10.2169/naika.97.2596 (2008).
- [18] 厚生労働省: 理学療法士・作業療法士の需給推計を踏まえた今後の方向性について <https://www.mhlw.go.jp/content/10801000/000499148.pdf> (2019).
- [19] 近藤国嗣: 回復期リハビリテーション医療—これまでの20年, これからの20年—, *The Japanese Journal of Rehabilitation Medicine*, Vol. 58, No. 5, pp. 468–481, DOI: 10.2490/jjrnc.58.468 (2021).
- [20] 島田洋一 (編), 高橋仁美 (編): 整形外科術後理学療法プログラム (第3版), メジカルビュー社 (2020).
- [21] 對馬均: 理学療法ハンドブック 第1巻 理学療法の基礎と評価, 協同医書出版社 (2000).
- [22] 徳田安春: 診断エラー学のすすめ, 医療の質・安全学会誌, Vol. 13, No. 1, pp. 53–56, DOI: 10.11397/jsqsh.13.53 (2018).
- [23] 奈良勲: 理学療法概論, 医歯薬出版株式会社 (2018).
- [24] 奈良勲 (監修): 基礎科学を融合した理学療法推論の実際, 運動と医学の出版社 (2022).
- [25] 日本運動器理学療法学会: 股関節機能障害理学療法ガイドライン, 理学療法ガイドライン第2版, 医学書院, pp. 659–665 (2021).
- [26] 日本整形外科学会/日本骨折治療学会 (監修): 大腿骨頸部/転子部骨折診療ガイドライン2021 (改訂第3版), 南江堂 (2021).
- [27] 平山朋子, 松下佳代, 西村敦: 医療教育における臨床推論を促進する「考える OSCE-R」の開発, 教師学研究, Vol. 18, pp. 23–33, DOI: 10.20590/jaehd.18.0_23 (2016).

- [28] 堀寛史: 理学療法推論とはなにか? —理学療法推論概論—, 理学療法推論, Vol. 1, No. 0, pp. 1–24, DOI: 10.82615/pmds.1.0.1 (2026).
- [29] マイケル ポランニー (著), 高橋 勇夫 (訳): 暗黙知の次元, ちくま文庫 (2003).
- [30] 宮本誠人, 松下光範, 高岡良行, 堀寛史: 理学療法初学者の支援を目的とした動作分析テキストの構造の可視化, 人工知能学会 (第 36 回) 全国大会論文集, 1I1OS604, DOI: 10.11517/pjsai.JSAI2022.0_1I1OS604 (2022).
- [31] 山路雄彦, 渡邊純, 浅川康吉, 松田祐一, 白田滋, 遠藤文雄, 内山靖, 坂本雅昭, 山口晴保, 中澤次夫, 茂原重雄: 理学療法教育における客観的臨床能力試験 (OSCE) の開発と試行, 理学療法学, Vol. 31, No. 6, pp. 348–358, DOI: 10.15063/rigaku.kj00003654240 (2004).
- [32] 吉田龍洋, 畠山駿弥, 堀寛史, 佐々木恭志郎, 松下光範: 急性期病院における理学療法士のリスク管理支援に関する一検討, 電子情報通信学会 HCG シンポジウム 2022 論文集, B-5-3 (2022).
- [33] 吉田龍洋, 畠山駿弥, 杉本明文, 堀寛史, 佐々木恭志郎, 高岡良行, 松下光範: 医療情報における着目点提示による理学療法初学者の臨床推論, 電子情報通信学会 HCG シンポジウム 2023 論文集, B-2-2 (2023).
- [34] 芳野純, 二渡玉江, 大谷健, 白田滋: 自立した理学療法士が獲得すべき能力に関する質的研究, 理学療法学, Vol. 37, No. 6, pp. 410–416 (2010).
- [35] 徳永剛, 今井孝樹, 流合慶多: 若手理学療法士の苦手としている臨床能力 及び学びたい臨床能力の調査 —主観的評価表を用いて—, 理学療法さが, Vol. 6, pp. 7–13 (2020).