

Ⅲ. 関連研究報告

プログラム評価におけるフィデリティ尺度の開発と妥当性の検証に関する 海外文献紹介

研究分担者：○吉田光爾¹⁾

研究協力者：片山優美子²⁾，下平美智代¹⁾

1) 独) 国立精神・神経医療研究センター 精神保健研究所 社会復帰研究部

2) 長野大学

要旨

目的：EBPの普及の上では、効能研究や効果研究によって有効性が明らかにされた心理社会的介入プログラムのモデルを同定し、その定式化・標準化を行いプログラムの実施・普及をはかることが重要である。フィデリティ尺度は、このプログラムの普及を円滑にするために、オリジナルに開発されたプログラムモデルに忠実に行われている程度（範囲）を評価する尺度である。こうしたフィデリティ尺度の開発や妥当性の方法論について、我が国では十分な文献・研究が導入されているとは言えない。本研究では Mowbray らの“Fidelity Criteria: Development, Measurement, and Validation(2003)”に基づき、フィデリティ開発の手順について簡略に紹介を行い日本におけるフィデリティ開発・検証研究の進展の一助とすることを目的とする。

概要：本論文はレビュー論文のスタイルをとり、健康や教育領域における論文から実例を引き、フィデリティ基準の開発、測定、および妥当化における道筋の要点を述べている。

結果：フィデリティ開発における3つの主要なステップは、①プログラムにおける有効な支援の構成要素を特定するフィデリティ尺度の開発、②指標を測定するための定量的なデータの収集・フィデリティの測定、③フィデリティ尺度の信頼性と妥当性の確認、である。③については(1)評価者間信頼性・再テスト信頼性、(2)確認的因子分析・内的一貫性、(3)他のプログラムとの比較、(4)収束的妥当性、(5)予測妥当性の検証などのアプローチを単一・または複数もちいて検証される。

考察：プログラムを対象としたフィデリティ尺度の開発には多くの時間がかかるため、1つの研究・論文でこの3ステップすべてを記述することは容易ではないが、その開発過程を①確立されたプログラムの有効な要素を忠実にフィデリティ尺度が反映しているか、②尺度が有効な要素を信頼性・妥当性をもって測定できているか、③その得点がアウトカムを予測できるか、という3つに分け検証するなど、研究を整理して記述するうえでは重要であろう。

い、それをもとにプログラムの実施・普及をは

A. 研究の背景

近年世界的にEBPに基づく医療・社会福祉実践に関する重要性が指摘されている。EBPの普及の上では、効能研究や効果研究によって有効性が明らかにされた心理社会的介入プログラムのモデルを同定し、その定式化・標準化を行

かることが重要である。モデルが明確に定式化・標準化されることで、そのプログラムの複製・普及が可能になるからである。モデルの定式化・標準化のためには、個別プログラムへのプロセス評価を実施して、効果的なプログラム

援助要素を抽出・指定することが必要となる。そのうえで検討された効果的なプログラム援助要素が、効果的普及モデルの実施マニュアルに記述されるわけである。

次に必要なのは、そうしたマニュアル等にそって実施されたプログラムが、どの程度の質を保って実施されているかを評価することである。この評価にはフィデリティ尺度 (fidelity scales) が用いられる。フィデリティ尺度は、あるプログラムが、オリジナルに開発されたプログラムモデルに忠実に実行されている程度 (範囲) を評価する尺度である。その評価項目は上記で抽出された効果的なプログラムの援助要素によって構成されており、この尺度を用いることで、各々の実践が定められた基準通りに実施・導入されているかという「度合い」を系統的に評価できるのである。フィデリティ尺度を用いることは以下の意義をもつ。まず各々の活動状況が明らかになると同時に、現場への具体的な改善や指針に関するフィードバックを通じて実践の質を改善する。また、行政機関に対してプロジェクトがどの程度目標を達成しているかという状況を報告することが可能になり、情報を公共に公開することも可能になる。またプログラムの効果は、プログラムの実施状況に大きく影響されるため、フィデリティ尺度による評価による履行状況のモニタリングをすることは、サービスの質の担保の面でも重要である。

しかし、こうしたフィデリティ尺度の開発や妥当性の方法論について、我が国では十分な文献・研究が導入されているとは言えない。本研究では 2003 年に発表された Mowbray らの “Fidelity Criteria: Development, Measurement, and Validation”¹⁾ に基づき、フィデリティ開発の手順について簡略に紹介を行い、日本におけるフィデリティ開発・検証研究の進展の一助とすることを目的とする。

B. 概要

1) 本論文の概要について

本論文はレビュー論文のスタイルをとり、1995 年から発表当時までにおいて、Psych Abstracts、ERIC、Social Science Index、Social Work Abstracts、MedSearch から検索された、健康や教育領域における論文から実例を引き、フィデリティ基準の開発、測定、および妥当化における道筋の要点を述べている。本報告ではフィデリティ開発の立場から、その概要について紹介する。

C. 結果

1) フィデリティとは何か？

フィデリティは「オリジナルに開発されたプロトコールに対する実際の治療実施の順守」として定義される (Orwin, 2000, p.S310)²⁾。典型的には、尺度はフィデリティ (実施されるプログラム化された臨床的介入が、そのベースとなる実証的に検証されたモデルに匹敵すること) を定量化するために開発される (Drake et al. 2001)³⁾。

2) なぜフィデリティを評価するのか？

フィデリティを評価する理由のとしてよくあげられるものは、ネガティブもしくは不明瞭な知見について説明する必要性からである (Hohmann & Shear, 2002)⁴⁾。すなわちあるプログラムが成功しなかった場合、意図したモデルに対するプログラムの順守の度合いを測定せねば、その原因がプログラムモデルそのものの失敗を反映したものなのか、意図したようにモデルを実践することに対して失敗したことによるのか、判定できない (Chen, 1990)⁵⁾。フィデリティ基準を確立することで、元のプログラムに対するアドヒアランスを測定可能にすることは、治療をより標準化し、プログラムを厳密に複製することを可能にする。またメタ分析のために、フィデリティ尺度を手にするには、各々の治療の意義ある比較を生み出す上で助けとなる (Banks, McHugo, Williams, Drake, & Shinn, 2001; Bond, Williams, Evans, et al. 2000)^{6),7)}。さらに、ある確立された支援モ

デルが、妥当な基準を用いて余所に複製されるときには、フィデリティ尺度で測定される点数は、アウトカムを予測することができる (Blackly et al, 1987; Paulson et al, 2002) 8),9)。

3) フィデリティ開発におけるステップ

研究者たち (McGrew et al, 1994; Teague, Bond, & Drake, 1998) 10),11)は、フィデリティ開発における3つの主要なステップについて記述している。すなわち①フィデリティ尺度の開発、②フィデリティの測定、③フィデリティ尺度の信頼性と妥当性の確認である。

①フィデリティ尺度の開発

第一に、フィデリティ尺度を開発しようとするプログラムにおける、有効な支援の構成要素を特定することが必要である。そして、それらが客観的で測定可能なように、それぞれの指標についてのデータの由来を記述し、評価尺度における特定のアンカーポイントを含め、指標もしくは候補となる構成要素のための操作的定義を設定する。

一般的に、これらを行うために、以下のような方法がとられている。

- (1) 効能、有効性が証明されているか、または少なくとも受容をとまなう特定のプログラムモデルからの描出
- (2) 専門家の意見を集める—専門家の調査、および文献レビュー
- (3) 質的調査—何が有効に働くかに関してのユーザーと支持者達の意見を収集する、さまざまなプログラムへのサイト訪問など

(ただし、基盤となるプログラムがしばしば RCT 等によって有効性が確認されているわけではないという問題がある)

②フィデリティ尺度の測定

第二のステップは、指標を測定するためにデータを集めることが必要である。

フィデリティのための定量的データを集めるもっとも一般的な方法は：

- (1) プロジェクト記録やクライアント記録、サイト観察、あるいは録画されたセッションに基づく専門家による評定、
- (2) サービスを提供する個人や受け取る個人による調査またはインタビュー、である。

ただし、これらには評価バイアスの問題がある。スタッフからの情報に関しては社会的望ましさのバイアスの問題、サービス利用者から評価してもらうことは自発的な参加者が多くなるというバイアスがあるため、これらのバイアスに対しては、各尺度ポイントを行動に基づく固定された基準 (アンカーポイント) にするなどの対応が考えられる。

また、プログラムの評価には《構造基準》と《プロセス基準》が存在する。プログラム構造の評価は職員水準や特徴、ケースロードの大きさ、予算、訴訟法、コンタクトの頻度や集中度などの測定を網羅する (Orwin, 2000) 2)が、これらの情報は、主観的判断を要求されることが少なく、多くの場合既存の記録物を通して取得することができる。一方《プロセス基準》は、プログラムの様式、スタッフ-クライアント関係、クライアント-クライアント関係、治療の個別化、あるいは情緒的雰囲気を含む。これらの基準と関連するプログラム業績の評定は、多くの場合、観察やインタビューや他のデータ源に基づくより主観的な判断を要求する。従って、より多くの時間と労力を必要とし、よりコストが高く、たとえ回答基準がうまく固定化されていたとしても信頼性はより低くなる傾向がある (Bond et al, 1997) 12)。これらの両側面の指標を用いることで、重要な特徴を測ることを可能にする。

③フィデリティ尺度の信頼性と妥当性の検証

第三のステップは、妥当性 (予測妥当性、判別妥当性、構成概念妥当性) と信頼性の観点から指標を検証することである (Moncher

&Prinz, 1991)¹³。これまでの研究ではフィデリティ尺度の開発研究では以下のうち1つまたは複数のアプローチを採用していることが多い。

(1)評価者間信頼性や再テスト信頼性の確認：

Henggeler et al. (2002)¹⁴は、セラピストに関する家族の複数回答による評定の再テスト信頼性を算出し、またセラピストとそのスーパーバイザーの間のレイティングの相関係数を算出している。Weisman et al. (2002)¹⁵はセラピストのビデオをみた場合の三人の評価者の級内相関係数を算出している。Clarke (1998)¹⁶は思春期の抑うつのための14の治療セッションに関してフィデリティアセスメントを行い、カッパ係数を算出している。

(2)データの内部構造について検証的な確認・期待される結果との関連の分析：

確認的因子分析や、クロンバックの α 係数を算出した内的一貫性の分析、またはクラスタ分析による検証である。Bond et al.(1997)¹², Clarke (1998)¹⁶, Lucca (2000)¹⁷, and McGrew et al. (1994)¹⁰らは内的一貫性を報告しており、他方で Henggeler et al. (2002)¹⁴は確認的因子分析と内的一貫性の分析を行っている。

(3)他のプログラムとの比較：

他の異なるプログラム間でフィデリティを評価してその差を検証する方法である。模範的な最上のプログラムと、通常のプログラムを比較したり(Hernandez et al, 2001)¹⁸、ACTプログラムと通常のケースマネジメントを比較したり (Teague et al, 1995, 1998)^{19,11}、様々な援助付き雇用を比較したり (Bond et al, 1997)¹²、クラブハウスと通常の職業リハビリテーションモデルを比較したり (Lucca,2000)¹⁷、援助付住居と比較プログラムを対比させたり(Rog & Randolph, 2002)²⁰、複合的なプログラムの状況と単一の介入を比較したり、あるいは濃淡のあるグループの参加

者を比較したり (Orwin, 2000)²¹するものである。

(4)収束的妥当性の検証：

プログラムに関する異なる2つのソースからの情報に関する一致を検証する方法でプログラムの収束的妥当性を検証する。例えば Blakely et al. (1987)⁸は、サイトにおける観察と、記録や文書からの情報を比較した。Macias et al. (2001)²¹は CRESS (Clubhouse research and Evaluation Screening Survey) 上の、クラブハウスの基準順守に関する自己評価と、実地の認証手続きの結果を、CRESSが認証した機関とそうでない機関について比較している。Lucca (2000)¹⁷はクラブハウスのフィデリティスコアと Principles of Psychosocial Rehabilitation scaleのスコアをつきあわせることで、収束的妥当性を検証している。

(5)予測妥当性の検証：

フィデリティ測定結果が、期待されるアウトカムと関連しているかどうかを検証してフィデリティの予測妥当性を確かめるものである。アプローチの例としては、援助付き雇用のフィデリティスコアが利用者のアウトカムと関連していることを明らかにした (Becker et al, 2001)²²。またACTのフィデリティモデルは明らかに入院の減少率と関連していた(McGrew et al, 1994)¹⁰。薬物と精神障害の併発に関するチームにおいては、レートが高いほど治療への残存率がよく、アルコールと薬物の使用現象に繋がり、寛解率も高かった (McHugo, Drake, Teague, & Xie, 1999)²³。また、メンタルヘルスに関する問題でリファーされた子供たちの割合とメンタルヘルスの支援を受けた高さが報告されている(Friesen et al., 2002)²⁴。さらに家族に焦点化したセラピーのフィデリティスコアと患者の寛解との関係が調べられている(Weisman et al., 2000)²⁵。

C. 考察

ここまでフィデリティ尺度の開発に関して整理してきた。フィデリティ開発には3つのステップがあり①フィデリティ尺度の開発（有効な支援要素の特定）、②フィデリティ尺度の測定、③フィデリティ尺度の信頼性・妥当性の検証、という3つのステップからなることが説明された。しかしフィデリティ尺度の開発研究は、いわゆる伝統的な心理・精神医学的な尺度に比べて歴史が浅く、その尺度の開発に関しては十分な研究蓄積がない段階といえるだろう。特に信頼性・妥当性の検証については必ずしも十分に定式化されているわけではなく、予測妥当性・判別妥当性・構成概念妥当性・信頼性の観点から、研究によって1つまたは複数のアプローチを採用して尺度の検証をしていることが多いようである。

これらの理由については、フィデリティ尺度が、個人の内面等ではなく、社会プログラムを対象にしているものと関連しているのではないかと考える。社会プログラムが成熟し、一定の成果を上げ、その成果が認知されるようになるまでには、場合によっては多くの時間を要する場合が少なくない。さらに本論文で紹介された「①成熟したプログラムの中から必要な構成要素を抽出する」には、さらに文献レビューや質的調査などの時間を要することになる。そのうえで上記②のステップでデータを実測し、さらに③のステップで信頼性・妥当性を検証する、とくにフィデリティ得点がアウトカムと関連するかという重要な観点である予測妥当性も含めた信頼性・妥当性を検証する場合には、プログラムが成果をあげるまでの時間をフォローする必要があり、それに要する時間は増大する。すなわちプログラムの成熟、フィデリティの①作成、②測定、③検証（一時点・長期予後）のステップごとに時間を要するため、1つのフィデリティ尺度の開発全体にはかなり長期的な関わりが必要になってくる。多くの論文が信頼性・妥当性の検証として必ずしもすべての側面をカバーしていないのは、1論文

では、この過程全体を記述しきれない、あるいは時間の経過を要するため、各論文が、その可能な範囲で信頼性・妥当性を検証しているともいえるだろう。

フィデリティ尺度は、まず有用性のあるプログラムが確立していることを前提にしている。そして確立されたプログラムを、フィデリティ尺度によって有効な要素に分解・構成することで複製を可能にすること、そのうえでそれを測定し、高い忠実度得点を維持することで元のプログラムと同等の高いアウトカムを達成することが期待されるわけである。そのように考えると、フィデリティ尺度の開発において重要な点は、Ⅰ：確立されたプログラムの有効な要素を忠実にフィデリティ尺度が反映しているか（紹介論文のステップ①）、Ⅱ：尺度が有効な要素を信頼性・妥当性をもって測定できているか（紹介論文のステップ②・③）、Ⅲ：その得点がアウトカムを予測できるか（紹介論文のステップ③）、という3つに整理したうえで、それを1つずつ検証していくなどの方法も、研究を整理して記述するうえでは重要であろう。

日本では現在、フィデリティ開発に関する研究が十分に蓄積されているとは言えないが、通常の尺度開発とは異なるこのような性格を踏まえ研究が重ねられることが重要であると考えられる。

D. まとめ

目的：EBPの普及の上では、効能研究や効果研究によって有効性が明らかにされた心理社会的介入プログラムのモデルを同定し、その定式化・標準化を行いプログラムの実施・普及をはかることが重要である。フィデリティ尺度は、このプログラムの普及を円滑にするために、オリジナルに開発されたプログラムモデルに忠実に実行されている程度（範囲）を評価する尺度である。こうしたフィデリティ尺度の開発や妥当性の方法論について、我が国では十分な文献・研究が導入されているとは言えない。本研究では Mowbray

らの“Fidelity Criteria: Development, Measurement, and Validation(2003)”に基づき、フィデリティ開発の手順について簡略に紹介を行い日本におけるフィデリティ開発・検証研究の進展の一助とすることを目的とする。

概要: 本論文はレビュー論文のスタイルをとり、健康や教育領域における論文から実例を引き、フィデリティ基準の開発、測定、および妥当化における道筋の要点を述べている。

結果: フィデリティ開発における3つの主要なステップは、①プログラムにおける有効な支援の構成要素を特定するフィデリティ尺度の開発、②指標を測定するための定量的なデータの収集・フィデリティの測定、③妥当性と信頼性フィデリティ尺度の信頼性と妥当性の確認、である。③については(1)評価者間信頼性・再テスト信頼性、(2)確認的因子分析・内的一貫性、(3)他のプログラムとの比較、(4)収束の妥当性、(5)予測妥当性の検証などのアプローチを単一・または複数もちいて検証される。

考察: プログラムを対象としたフィデリティ尺度の開発には多くの時間がかかるため、1つの研究・論文でこの3ステップすべてを記述することは容易ではないが、その開発過程を①確立されたプログラムの有効な要素を忠実にフィデリティ尺度が反映しているか、②尺度が有効な要素を信頼性・妥当性をもって測定できているか、③その得点がアウトカムを予測できるか、という3つに分け検証するなど、研究を整理して記述するうえでは重要であろう。

E. 健康危険情報 なし

F. 研究発表 なし

G. 知的財産権の出願・登録状況

1. 特許取得 なし
2. 実用新案登録 なし
3. その他 なし

文献

- 1) Mowbray, K, Holter, M. C., Teague G. B. & Bybee, D.(2003). Fidelity Criteria: Development, Measurement, and Validation”, *American Journal of Evaluation* 24(3), 315-340.
- 2) Orwin, R. G. (2000). Assessing program fidelity in substance abuse health services research. *Addiction*,95(Suppl.3), S309-S327.
- 3) Drake, R., Goldman, H., Leff, H., Lehman, A., Dixon, L., Mueser, K.,&Torrey,W. (2001).Implementing evidence-based practices in routine mental health service settings. *Psychiatric Services*, 52, 179-182.
- 4) Hohmann, A. A., & Shear, M. K. (2002). Community-based intervention research: Coping with the“noise” of real life in study design. *American Journal of Psychiatry*, 159,201-207.
- 5) Chen,H.(1990). *Theory-driven evaluations*. Thousand Oaks, CA: Sage.
- 6) Banks, S., McHugo, G. J., Williams, V., Drake, R. E., & Shinn, M. (2001). A prospective meta-analytic approach in a multi-site study of homelessness prevention. In J. M. Herrell & R. B. Straw (Eds.),
- 7) Bond, G. R., Evans, L., Salyers, M. P., Williams, J., & Kim, H. W. (2000). Measurement of fidelity in psychiatric rehabilitation. *Mental Health Services Research*, 2(2), 75-87.
- 8) Blakely, C. H., Mayer, J. P., Gottschalk, R. G., et al. (1987). The fidelity-adaptation debate: Implications for the implementation of public sector social programs. *American Journal of*

- Community Psychology, 15, 253-268.
- 9) Paulson, R. I., Post, R. L., Herinckx, H. A., & Risser, P. (2002). Beyond components: Using fidelity scales to measure and assure choice in program implementation and quality assurance. *Community Mental Health Journal*, 38, 119-128.
 - 10) McGrew, J. H., Bond, G. R., Dietzen, L., & Salyers, M. (1994). Measuring the fidelity of implementation of a mental health program model. *Journal of Consulting and Clinical Psychology*, 62, 670-678.
 - 11) Teague, G. B., Bond, G. R., & Drake, R. E. (1998). Program fidelity and Assertive Community Treatment: Development and use of a measure. *American Journal of Orthopsychiatry*, 68, 216-232.
 - 12) Bond, G. R., Becker, D. R., Drake, R. E., et al. (1997). A fidelity scale for the individual placement and support model of supported employment. *Rehabilitation Counseling Bulletin*, 40, 265-284.
 - 13) Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review*, 11, 247-266.
 - 14) Henggeler, S.W., Schoenwald, S. K., Liao, J. G., Letourneau, E. J., & Edwards, D. L. (2002). Transporting efficacious treatments to field settings: The link between supervisory practices and therapist fidelity in MST programs. *Journal of Clinical Child and Adolescent Psychology*, 31, 155-167.
 - 15) Weisman, A., Tompson, M. C., Okazaki, S., Gregory, J., et al. (2002). Clinicians' fidelity to a manual-based family treatment as a predictor of the one-year course of bipolar disorder. *Family Process*, 41, 123-131.
 - 16) Clarke, G. (1998). Intervention fidelity in the psychosocial prevention and treatment of adolescent depression. *Journal of Prevention and Intervention in the Community*, 17, 19-33.
 - 17) Lucca, A. M. (2000). A Clubhouse fidelity index: Preliminary reliability and validity results. *Mental Health Services Research*, 2, 89-94.
 - 18) Hernandez, M., Gomez, A., Lipien, L., Greenbaum, P. E., et al. (2001). Use of the system-of-care practice review in the national evaluation: Evaluating the fidelity of practice to system-of-care principles. *Journal of Emotional and Behavioral Disorders*, 9, 43-52.
 - 19) Teague, G. B., Drake, R. E., & Ackerson, T. H. (1995). Evaluating use of continuous treatment teams for persons with mental illness and substance abuse. *Psychiatric Services*, 46, 689-695.
 - 20) Rog, D. J., & Randolph, F. L. (2002). A multi-site evaluation of supported housing: Lessons learned from cross-site collaboration. In J. M. Herrell & R. B. Straw (Eds.), *Conducting multiple site evaluations in real-world settings*. *New Directions for Evaluation*, no. 94 (pp. 61-72). San Francisco, CA: Jossey-Bass.
 - 21) Macias, C., Propst, R., Rodican, C., & Boyd, J. (2001). Strategic planning for ICCD clubhouse implementation: Development of the Clubhouse Research and Evaluation Screening Survey (CRESS). *Mental Health Services Research*, 3, 155-167.
 - 22) Becker, D. R., Smith, J., Tanzman, B., Drake, R. E., & Tremblay, T. (2001). Fidelity of supported employment programs and employment outcomes. *Psychiatric Services*, 52, 834-836.
 - 23) McHugo, G. J., Drake, R. E., Teague, G. B., & Xie, H. (1999). The relationship between model fidelity and client outcomes in the New Hampshire Dual

Disorders Study. *Psychiatric Services*, 50, 818-824.

24) Friesen, B. J., Green, B. L., Kruzich, J. M., Simpson, J., et al. (2002). Guidance for program design: Addressing the mental health needs of young children and their families in early childhood education settings. Retrieved May 23, 2002, from Portland State University, Research & Training Center on Family Support and Children's Mental Health Web site: <http://www.rtc.pdx.edu/pgProjGuidance.php>.

25) Weisman, A., Nuechterlein, K. H., Goldstein, M. J., & Snyder, K. S. (2000). Controllability perceptions and reactions to symptoms of schizophrenia: A within-family comparison of relatives with high and low expressed emotion. *Journal of Abnormal Psychology*, 109, 167-171.

日本版 IPS 型就労支援のフィデリティ評価ツール開発に係る研究

研究分担者：○下平美智代¹⁾

研究協力者：山口創生¹⁾、吉田光爾¹⁾、佐藤さやか¹⁾、市川健¹⁾、古家美穂¹⁾

種田綾乃¹⁾、片山優美子²⁾、小川友季³⁾、伊藤順一郎¹⁾

1) 独) 国立精神・神経医療研究センター 精神保健研究所 社会復帰研究部

2) 長野大学

3) 独) 国立国際医療研究センター 国府台病院

要旨

本研究では、日本版 IPS 型就労支援フィデリティ評価ツールを開発し、ツールの信頼性と妥当性の検証を行うことを第一の目的とした。さらに、日本の IPS 型支援実施機関の特徴を記述することを副次的目的とした。

10 項目の GOI、25 項目の JIPS-25 を作成した。この評価ツールを用いて、日本で精神障害者を対象とした個別就労支援を実施している 17 機関を対象にフィデリティ調査を行った。結果として、1 機関を除く 16 機関の JIPS-25 評価点が 74 点以上であった (IPS-25 では、73 点以下について「援助付雇用とはいえない」と判定する)。

調査で取得したデータから、GOI、JIPS-25 共に評価者間信頼性が確認された。妥当性については、IPS 標榜群 (n=12) はそうでない群 (n=5) よりも GOI および JIPS-25 の得点が高い傾向にあったことから、弁別的妥当性が確認できた。また、IPS を標榜する機関はそうでない機関と比較して、新規登録者数は少ない傾向にあり、就職率および離職率は高い傾向にあった。また、JIPS-25 と就職率および離職率とは有意な正の相関が示された。ただし、JIPS-25 の下位項目「地域ベースのサービス」の評価の高い機関ほど離職率は低いという結果も示された。

A. 研究の背景

重度精神障害者の個別援助付雇用システムとしてアメリカで開発された IPS (Individual Placement and Support) 型就労支援は、複数の無作為化比較対照試験 (RCT) により、準備訓練型の支援よりも有効であることが実証されている¹⁾。そのため、IPS 型就労支援プログラムは EBP (Evidence based practice: 根拠に基づく実践) であることが国際的に知られている。我が国では、2005 年前後から IPS 実践が複数の実践機関もしくは実践家に支援技法として取り入れられてきたが、必ずしも

オリジナルモデルに忠実なプログラムとして提供されてきたわけではなかった⁴⁾。そのため、IPS 型就労支援をよく理解し、効果的支援を提供していた実践家はその機関を去ると、後には継承されていかないという状況もみられた。

その機関でオリジナルモデルに忠実な支援プログラムが提供されているかどうかについては、フィデリティ評価によって確認することができるが、本研究班発足当時 (2011 年 4 月)、日本版のフィデリティ尺度は存在していなかった。フィデリティとは、「ある介入がそ

のプログラムもしくはオリジナルに開発されたプログラムモデルに忠実に実行されている程度」³⁾と定義される。

本研究班では、平成 23 年度（2011 年度）に、オリジナルモデルの確認として、アメリカバーモント州におけるオリジナルモデルの視察と研修を実施した。さらに、IPS-25²⁾を参照し、平成 24 年度（2012 年度）に研究者、有識者、および実践家から成るワーキンググループを結成して「日本版 IPS 型就労支援標準モデル」を作成した⁵⁾。

本研究では、この標準モデルと IPS-25 を基に、日本版 IPS 型就労支援フィデリティ評価ツールを開発し、日本で精神障害者を対象とした個別就労支援を実施している機関を対象に実地試用を行い、ツールの信頼性と妥当性の検証を行うことを第一の目的とした。さらに、日本の IPS 実施機関の特徴を記述することを副次的目的とした。

B. 方法

1. 日本版 IPS 型就労支援フィデリティ評価ツールの開発

日本版 IPS 型就労支援標準モデル（別添）と IPS-25 を基に、まずはフィデリティ評価項目のたたき台を作成し、標準モデルのワーキンググループに入っていた研究者らが、フィデリティ評価ツール開発メンバーとなり、項目を一つずつ確認し、文章を修正する作業を行った。また、開発メンバーからは、フィデリティ評価を補うものとして、全般的な組織体制が評価できる指標の作成が提案され、GOI（General Organization Index：一般組織構造指標）も同時期に作成された。

2. ツールの試用および信頼性と妥当性の検証

①対象

2012年8月に当研究班によるアンケート調査⁴⁾対象となった、IPS を標榜もしくは要素

を取り入れているとする 21 機関、その後、リストに加えられた 3 機関、および IPS を標榜していないが、個別ニーズベースの就労支援に取り組む 2 機関の合計 26 機関を対象とした。その内、訪問調査を承諾した 17 機関にフィデリティ調査を実施した。

17 機関の基礎属性を表 1 に示した。就労移行支援事業所が最も多く 11 機関（65%）、次いで医療機関 4 機関（24%）、生活訓練 1 機関（6%）、ACT を専門に行っている訪問看護ステーション 1 機関（6%）であった。これら 17 機関の内、IPS を標榜している機関が 12 機関、部分的に IPS を取り入れているとする機関が 3 機関、IPS ではないが個別就労支援を行っている機関が 2 機関であった。

②手続き

2 名の調査員が各機関に訪問し、本研究で開発された評価票（後述）に沿って調査を実施した。調査所用時間は 1 機関あたりおよそ 6 時間であった。評価方法は、当該機関の管理者および就労支援スタッフ（ES）へのインタビュー、およびスタッフが日々つけている利用者別の記録の調査であった。その他参考として実際の支援に調査員が同伴して支援の様子を観察した。調査実施期間は 2013 年 9 月 20 日から 2014 年 2 月 4 日までだった。

③分析

ツールの評価者間信頼性の分析には Pearson の積率相関係数の算出を行った。また、弁別的妥当性の検証のために、IPS を標榜する群とそうでない群（部分的に IPS を取り入れているもしくは IPS ではないと言明している群）とに分け、フィデリティ得点を Student の t 検定および Mann-Whitney の U 検定で比較した。また、JIPS-25 と利用者の就職率および離職率の関連を Pearson の積率相関係数を算出することで検討した。

④倫理的配慮

本研究は、国立精神・神経医療研究センターの倫理審査委員会にて審議にかけられ、認可を受けて実施した。

C. 結果

1. 評価ツールの開発

ツール開発作業に要した期間は、2013年4月1日～8月31日までの5ヵ月間で、その内、開発メンバーが議論したのは、1回3時間の合計5回であった。この間、評価ツールの調査票としての形成および評価の際の計算シートを作成が行われた。

前述のように、開発メンバーからは、フィデリティ評価を補うものとして、全般的な組織体制が評価できる指標の作成が提案され、GOI (General Organization Index : 一般組織構造指標) も同時期に作成された。

結果として、10項目のGOIと25項目のフィデリティ評価ツールが開発され、この25項目版のツールはJIPS-25と名付けられた。完成したJIPS-25第1版を付録1に、GOI第1版を付録2に収録している。

GOIは10項目の各項目について、「いる」「いない」もしくは「ある」「ない」の2者択一で評価するようになっている。「いる」もしくは「ある」の回答で1点ずつ加算し、0点から10点までで評価する。点数が高いほど一般組織構造がIPS型就労支援プログラムを実施する体制として評価が高いと判断される。

JIPS-25は、各項目について、5段階で評定するようになっており、各段階には明確なアンカーポイントが付されている。また、IPS-25に合わせて、「スタッフ配置」3項目、「組織」8項目、「サービス」14項目から成っている。なお、「サービス」には日本版標準モデルに合わせた2項目が配点評価しない参考項目として掲載されている。JIPS-25は全ての項目の点数を合算して合計点を出す場合は、

最低25点、最高で125点となる。下位尺度ごとに合計点を出す場合は、「スタッフ配置」3点から15点、「組織」8点から40点、「サービス」14点から70点の範囲となる。

JIPS-25も点数が高いほどフィデリティが高いと判定される。

なお、フィデリティ評価ツールは、内的一貫性を重要視する心理測定的尺度とは異なり、そのプログラムが本来目指すものに近いかどうかを評価するため、3つの下位尺度は因子分析等統計学的分析によって分類されたものではなく、エキスパートコンセンサスにより作成され分類された項目となっている。

2. ツールの信頼性と妥当性

①評価者間信頼性

フィデリティ調査には2名の調査員が出向き、それぞれがフィデリティ評価ツールに沿って評価点をつけた。GOI、JIPS-25共に、項目ごとに二者の評価点について相関係数を算出した。GOIについては、完全に二者間が一致しており、相関係数を算出するまでもなかった。JIPS-25についての結果を表2に示した。結果として1項目を除きすべての項目について有意な高い正の相関が示された。「サービス」の項目1のみ、相関係数も低めで有意ではなかった。これは、「社会保障に関する個別相談を提供する」という項目であった。この項目については、二者の評価点を精査したところ、一方が4点もしくは5点としている場合にもう一方が5点もしくは4点と評価しているという不一致が起きており、3点以下の評点のときには不一致がなかった。このため、5段階評定のアンカーポイントの4点と5点の文章について確認を行ったところ、「仕事を始める前に利用者が、社会保障に関する専門知識のあるスタッフまたは外部の専門家により、包括的で個別化された社会保障に関するカウンセリングを受けられるよう、就労支援スペシャリストもしくはケアマネー

ジャーはコーディネートする。」という文章までは両者は同じで、評点 5 では、「彼らは利用者が勤務時間や給与の変化に関して意思決定をする必要がある際にも社会保障に関するカウンセリングを受けられるようコーディネートする。」という文章が追加されている。この項目は記録からではなく、管理者と ES にそれぞれインタビューを行い評価する項目であった。

② 弁別的妥当性

弁別的妥当性の検討のために、IPS を標榜する機関 (n=12) とそうでない機関および部分的に IPS を取り入れているとした機関

(n=5) のフィデリティ得点を比較した。表 3 に示したとおり、フィデリティ得点および下位尺度得点全てにおいて、IPS 群の方がそうでない群より高かった。統計学的な有意差を Student の t 検定および Mann-Whitney の U 検定で確認したところ、合計点および「組織」点で有意傾向のある差がみられた。

3. 対象機関のフィデリティと利用者登録数および利用者の就職実績

17 機関のフィデリティ評価の結果を表 4 に示した。GOI の平均は 6.8 点 (sd=1.4, 範囲 3-9) で、JIPS-25 の平均は 91.3 点 (sd=10.2, 範囲 68-104) であった。JIPS-25 の下位尺度「スタッフ配置」の平均は 9.1 点 (sd=3.0, 範囲 4-13)、「組織」の平均は 23.3 点 (sd=5.5, 範囲 16-33)、「サービス」の平均は 54.2 点 (sd=5.2, 範囲 44-61) であった。

17 機関中、前年度 (2012 年度) の実績のあった 16 機関の利用者数の平均および就職実績を表 5 に示した (1 機関は新規だった)。

登録者数は 10 名から 126 名と幅があり、平均 44.1 (sd=29.2) であった。男性 (平均 28.8 人, sd=16.8) の方が女性 (平均 15.4, sd=13.1) よりも全体的に登録者数が多かった。本研究の定義する一般雇用とその他を含めた

全体の就職率は 44.4% (sd=20.1, 範囲 8.8-77.7) であり、一般雇用の就職率が 35.5% (sd=18.9, 範囲 7.4-65.5) であった。なお、「一般雇用」の定義は、「一般求人および障害者求人による一般企業等における雇用である。ただし、障害者だけを一つのセクションに集めたような職場は含めない。」というものであった。

筆者らは、IPS 支援では、本人の表明があってからの「迅速な求職活動」が原則なため、就職しても長続きしないのではないかと、という批判的な疑問の声をよくきいた。そのため、本研究では暫定的な離職率を、前年度就職した者で翌年の 8 月現在退職している者の数を各機関にたずね、その数値を使って算出した。結果として、離職率の平均は 20.2%、標準偏差 17.2 で、全く離職者のなかった機関もあれば最大値で 66.7%の機関まであり、機関によって大きなばらつきがあった。

4. フィデリティと就職率および離職率

JIPS-25 と就職率について積率相関係数を算出したところ、有意な正の相関が示された ($r=0.56, p<.05$)。下位尺度項目については、「スタッフ配置」の項目 3:「就労支援スペシャリストの提供するサービス」は一般雇用での就職率と有意な正の相関が示された

($r=0.59, p<.05$)。「組織」の項目 3:「就労支援スペシャリストと職業リハビリテーションにおける支援スタッフとの連携」は就職率と有意傾向のある正の相関が示された

($r=0.49, p<.1$)。「サービス」の項目 3:「職業的アセスメント」は就職率 ($r=0.84, p<.05$) および一般雇用での就職率 ($r=0.65, p<.05$) とそれぞれ有意な正の相関が示された。同じく「サービス」の項目 5:「個別化された求職活動」は一般雇用での就職率と有意傾向のある正の相関が示された ($r=0.50, p<.1$)。一方、「サービス」の項目 8:「職場開拓の多様性」は一般雇用での就職率と有意傾向のある負の

相関が示された ($r=-0.52, p<.1$)。

JIPS-25 は離職率とも有意な正の相関が示された ($0.58, p<.05$)。下位尺度項目では、「組織」の「就労支援ユニット」($r=0.71, p<.01$)、「除外基準なし (1 ヶ月以内の求職活動)」($r=0.68, p<.05$)、「機関が一丸となって利用者の一般雇用実現に取り組む」($r=0.61, p<.05$) が離職率と有意な正の相関が示された。「サービス」の項目 3:「職業的アセスメント」($r=0.50, p<.1$)、項目 4「一般雇用のための迅速な求職活動」($r=0.48, p<.1$) とは有意傾向のある正の相関が示された。一方、「サービス」の項目 13:「地域ベースのサービス」と離職率は有意な負の相関 ($r=-0.66, p<.05$) が示された。

5. IPS を標榜する機関とそうでない機関との実績の比較

IPS を標榜する機関の特徴を知るために、利用者の登録実績および就職率、離職率を比較した。結果を表 6 に示した。登録者数は、IPS 標榜群が平均 40.3 ($sd=23.3$)、非 IPS および部分的 IPS 群は平均 52.6 ($sd=41.4$) で、IPS 群の方が全体的に登録者数が少ない傾向が見られた。ただし、どちらの群も標準偏差が高く、統計学的検定の結果、有意ではなかった。就職率については、IPS 標榜群で平均 48.4% ($sd=19.2$)、非 IPS および部分的 IPS 群で 35.5% ($sd=21.3$) であり、やはり平均値では IPS 群の方が高いように見えるが、標準偏差が大きく統計学的検定の結果、有意差は示されなかった。離職率については、IPS 標榜群で 25.8% ($sd=17.5$)、非 IPS 群および部分的 IPS 群は 8.4% ($sd=2.1$) で、その差は統計学的に有意であった。

D. 考察

日本版 IPS 型就労支援フィデリティ評価ツールとして、10 項目の GOI、25 項目のフィデリティ評価ツール JIPS-25 を作成した。

GOI、JIPS-25 共に評価者間信頼性が確認された。JIPS-25 の下位尺度「サービス」の項目 1 のみ、二者間の評価の一致度が低かった。このため、二者の評価点を精査したところ、4 点と 5 点で評価が分かれているケースが多く、3 点以下の評価点は一致が見られていた。この項目は管理者と ES にインタビューをし、その結果から評価する。4 点と 5 点の評価のための文章の差異について、インタビュー結果の解釈が評価者間で割れやすかったのであろう。この項目は、評価者が客観的評価をしやすいうように修正する必要があると考えられる。

妥当性については、IPS 標榜群はそうでない群よりも GOI および JIPS-25 の得点が高い傾向にあったことから、弁別的妥当性が確認できたと考えられる。

調査対象機関によって、利用登録者数や就職者数にはばらつきがあったが、それは、機関が就労支援を専門に提供しているサービス機関 (就労移行支援事業所など) か、ケアや ACT など就労支援をサービスの一部として実施しているかによって異なっていたと思われる。しかし、全体的に就職率および一般就労率は高いように考えられるが、このことについては、別途、準備訓練型の支援を行っている機関との比較が必要である。

対象機関 17 機関の JIPS-25 の平均点は 91 ($sd=10.2$) で、68 点から 104 点の幅があった。IPS-25 の評価では、73 点以下は「援助付雇用とはいえない」とされ、74 点から 99 点は「まずまずのフィデリティ」、100 点から 114 点は「良好なフィデリティ」、115 点から 125 点が「模範的フィデリティ」と判定される。73 点以下だったのは、68 点の 1 機関のみであるが、ここでは記録の不備により全体的に評価点が低かったということがある。またこの 1 機関は部分的に IPS を取り入れている機関の 1 つであった。IPS を標榜していない機関でも個別ニーズベースの就労支援に取

り組んでいる機関が今回は調査対象となったため、全体的にフィデリティは高かったといえるだろう。

IPS のフィデリティの高さと利用者の高い就職率とは関連することが先行研究で示されている¹⁾。本調査結果では、JIPS-25 の評価点は就職率とも離職率とも有意な正の相関を示した。今回は暫定的な就職率の計算方法であり、経時的なアウトカムではないため、断言はできないものの、可能性として IPS 型就労支援を採用すれば、利用者の就職はより達成されやすくなることが示唆された。一方で、フィデリティ得点と離職率についても有意な正の相関がみられたことから、IPS らしい支援をしていると、利用者は離職しやすいということが示唆されているようにも思われる。しかし、今回の離職率の計算は暫定的なものであるため、就労継続日数を追跡した調査結果をみて判断しなければならないだろう。一方で、「地域ベースのサービス」の評価の高かった機関は、離職率が低かったことから、利用者の就職後の支援機関外（地域や就職先）での支援の重要性が改めて強調されたといえるだろう。

IPS を標榜する機関はそうでない機関と比較して、登録者数が少ない傾向にあり、就職率および離職率は高い傾向にあった。IPS 型就労支援はその特徴として、少ないケースロード(ES1 に対してクライアント 20 人以下)、除外基準をもたない、本人の希望による迅速な求職活動がなされることを是としている。このため、IPS 標榜機関はたしかに IPS らしい支援に取り組んでいると考えられる。一方で、IPS 標榜機関の方が利用者の離職率が高いという結果も示された。これは、短期雇用を就職活動の一環として積極的に利用している可能性と、支援者ではなく利用者本人が主体的に選択している結果とも考えられる。保護的な機関では、利用者が就職したいという意志表明をしても、時期尚早であるとして引

き留めたり、その意を削ぐようにわざと働きかけたりする機関もあると聞く。本人の希望により就職し、その雇用が継続されるのが一番望ましいが、離職や転職は必ずしもマイナスの経験とはいえない。IPS 支援を利用する人にはチャレンジする権利が保障される一方でリスクをも自分で引き受けるということであり、ES はそのリスクを引き受けた利用者を集团的に後方で支援するというのが IPS のあり方ではないか。繰り返しになるが、今回の結果は、利用者の就労継続日数を追跡調査したものではないため、IPS 型就労支援と離職率あるいは就労継続日数との関係は別の調査で確認する必要がある。さらに、本研究では「地域ベースのサービス」の評価の高い機関ほど離職率が低いという結果も示された。

IPS 型就労支援プログラムでは、ES が支援機関の外に出て、地域あるいは利用者の勤務先で支援すること、特に利用者の就職直後の集中的な支援が展開されることが強調されている。本調査結果は、その重要性が改めて示唆される結果であったと考えられる。

E. 結論

日本版 IPS 型就労支援フィデリティ評価ツールの GOI と JIPS-25 を開発し、一定の信頼性と妥当性が確認された。JIPS-25 で評価者間信頼性の低かった 1 項目のみ、修正点として反映したい。IPS 標榜機関の特徴として、そうではない機関と比較して、新規登録者数が少ない傾向にあり、就職率や離職率が高い傾向にあった。ただし、「地域ベースのサービス」の評価点が高い機関ほど離職率は低いという結果が示された。

F. 健康危険情報 なし

G. 研究発表

1. 論文発表 なし

・種田綾乃, 山口創生, 佐藤さやか, 片山優

美子, 伊藤順一郎: 重度精神障害者に対する就労支援: individual placement and support を中心に. 精神保健研究, 27(60)73-79, 2014.

2. 学会発表

- ・下平美智代, 山口創生, 片山(高原)優美, 吉田光爾, 佐藤さやか, 伊藤順一郎: 「日本版 IPS 型就労支援スタンダード」(標準モデル)作成に係る研究. 日本精神障害者リハビリテーション学会 第 21 回沖縄大会. 沖縄, 2013.11.29.

H. 知的財産権の出願・登録状況

1. 特許取得 なし
2. 実用新案登録 なし
3. その他 なし

文献

- 1) Bond GR, Drake RE, Becker DR. An Update on Randomized Controlled Trials of Evidence Based Supported Employment. *Psychiatric Rehabilitation Journal* 31(4): 280-290, 2008.
- 2) Bond GR, Peterson AE, Becker DR, Drake RE: Validation of the Revised Individual Placement and Support Fidelity Scale (IPS-25). *Psychiatric Service* 63(8): 758-763, 2012.
- 3) Mowbray CT, Holter MC, Teague GB, Bybee D: Fidelity Criteria: Development, Measurement, and Validation. *American Journal of Evaluation*, 24(3): 315-340, 2003.
- 4) 下平美智代, 片山(高原)優美子, 山口創生, 小川友季: 日本の IPS 型就労支援実施機関を対象とした郵送調査結果の報告. 厚生労働科学研究費補助金 難病・がん等の疾患分野の医療の実用化研究事業(精神疾患関係研究分野)「「地域生活中心」を推進する、地域精神科医療モデル作りとその効果検証に関する研究」平成 24 年度総括・研究分担報告書(研究代表者 伊藤順一郎). pp223-230, 2013.
- 5) 下平美智代, 山口創生, 片山(高原)優美子, 吉田光爾, 佐藤さやか, 種田綾乃, 市川健, 伊藤順一郎, 日本版 IPS 型就労支援スタンダード作成委員会: 「日本版 IPS 型就労支援スタンダード(標準モデル)」作成に係る研究. 厚生労働科学研究費補助金 難病・がん等の疾患分野の医療の実用化研究事業(精神疾患関係研究分野)「「地域生活中心」を推進する、地域精神科医療モデル作りとその効果検証に関する研究」平成 24 年度総括・研究分担報告書(研究代表者 伊藤順一郎). pp231-244, 2013.

	就労移行	病院	生活訓練	ACT 訪看)
調査参加総数	11	4	1	1
割合 (%)	65	24	6	6
PS	7	3	1	1
部分PS	2	1	0	0
PSではない	2	0	0	0

項目番号	スタッフ配置	組織	サービス
	相関係数 (r)		
1	1.00	0.99	0.45 (ns)
2	1.00	0.78	0.94
3	0.81	0.85	0.86
4		0.97	1.00
5		0.93	0.82
6		0.81	0.99
7		0.79	0.93
8		0.84	1.00
9			1.00
10			1.00
11			0.77
12			0.80
13			1.00
14			0.92

Pearsonの積率相関分析 $p<0.01$ (両側検定)

	PS 標榜 n=12		PS ではない/部分的 n=5		Student's t test		Mann-Whitney's test
	平均値	標準偏差	平均値	標準偏差	t値	有意確率	漸近有意確率
GOI	7.0	1.2	6.2	1.9	1.0	0.311	0.514
Fidelity	94.2	9.0	84.4	10.3	2.0	0.069	0.091
スタッフ配置	9.3	2.9	8.8	3.5	0.3	0.788	0.789
組織	24.8	5.4	19.6	3.9	1.9	0.071	0.080
サービス	55.1	5.3	52.0	4.6	1.1	0.277	0.245

	平均値	標準偏差	最小値	最大値	(満点)
GOI	6.8	1.4	3	9	10項目 (10)
Fidelity	91.3	10.2	68	104	25項目 (125)
スタッフ配置	9.1	3.0	4	13	3項目 (15)
組織	23.3	5.5	16	33	8項目 (40)
サービス	54.2	5.2	44	61	14項目 (70)

	平均	標準偏差	最小値	最大値
2012年度登録者数	44.1	29.2	10	126
男性 数)	28.8	16.8	5	70
女性 数)	15.4	13.1	2	56
就職者数	21.8	20.9	3	76
「一般雇用」就職数	17.8	18.7	1	67
2013年8月現在退職者数	4.6	5.9	0	24
就職率(%)	44.4	20.1	8.8	77.6
一般雇用への就職率(%)	35.5	18.9	7.4	65.5
離職率(%)	20.2	17.2	0	66.7
「一般雇用」の定義：一般求人および障害者求人による一般企業等での雇用 障害者だけを一つのセクションに集めた職場は含めない)				
就職率＝就職者数／登録者数×100				
離職率＝2013年8月現在退職者数／就職者数×100				

	PS標榜 n=12		PSではない／部分的 n=5		<i>Student's t test</i>		<i>Mann-Whitney's test</i>
	平均値	標準偏差	平均値	標準偏差	t値	有意確率	漸近有意確率
2012年度登録者数	40.3	23.3	52.6	41.4	-0.8	0.453	0.610
就職率	48.4	19.2	35.5	21.3	1.2	0.248	0.193
一般雇用率	37.6	19.3	31.0	19.3	0.6	0.537	0.777
離職率	25.8	17.5	7.9	8.4	2.1	0.050	0.035