

支援機器実証試験における科学性

Scientific Aspects of Clinical Test of Assistive Products

○山内繁（特定非営利活動法人支援技術開発機構）

加藤誠志 諏訪基 中山剛 外山滋（国立障害者リハビリテーションセンター研究所）

Shigeru YAMAUCHI, Assistive Technology Development Organization

Seishi KATO and Motoi SUWA, Tsuyoshi NAKAYAMA, Shigeru TOYAMA, Research Institute, National Rehabilitation Center for Persons with Disabilities

Abstract: Scientific clinical test is indispensable for development of assistive products. This paper aims at producing protocols with objective evidence for evaluation and with clearly defined goals. Research design to be applied to the clinical test of assistive products is discussed to ensure robust evidence. The phases of clinical test help to define clear goals. Definition of phases to be applied specifically to the development of assistive products is proposed. These factors are examined in the papers of RESJA symposiums. Discussion is also made in relation to IRB review.

Key Words: Assistive products, IRB review, Protocols

1. はじめに

支援機器の実証試験の目的は、開発された機器が所期の機能を有することを検証することならびに適応、適合のための詳細な情報を得ることにある。当然、安全性の確認も必要であるが、安全性そのものは実証試験に先立つリスクアセスメントにおいて確立しているべきであって、実証試験によって安全性を確認すべきではない。

現在我が国で行われている支援機器実証試験においては主に2つの点で欠陥があるとの印象がある。その第一は実証試験における「検証」が主として（被験者または開発者による）主観評価によっており、客観的なエビデンスに基づいた検証がなされていない点である。第二は、実証試験が漫然と計画され、明確なゴール設定のないプロトコルがまかり通っている点にある。本報告ではこのような問題を解決するためのツールを探るとともに、現状の問題点を検証することにしたい。

2. 研究デザイン

人を対象とする研究においては、物理学、化学を基礎とした通常の工学におけると同様な決定論的アプローチを取ることができない。それは、介入に対して人体の反応を支配する要因が多すぎてそのすべてを制御することは不可能であることに起因する。

このために、生物医学領域において前世紀半ばから発展したのが推測統計学に基づいた仮説検証型の研究デザインである。研究デザインの詳細は教科書^(1,2)を参考にしたいが、簡単に概観しておく。研究デザインは観察研究と介入研究とに分けられる。観察研究は、現象を在るがままに観察する研究であって、日常的に使用している支援機器の使用状況を観察する研究である。観察研究でよく用いられる研究デザインはコホート研究と横断研究である。前者はあるグループの挙動を追跡観察するのに対し、後者はある時点における挙動を記録する。

介入研究は、被験者を2群に分けて反応を比較する場合、通常の支援を超える支援(医学の場合は診療)に対する反応を記録する場合がある。介入研究は比較対照研究と対照のない研究に分けられる。対照のない研究は介入に対するアウトカムを記録するだけであるが、比較対照研究は2群の

対象者のアウトカムを比較する。

現在最も強力な研究デザインとされているのはランダム化臨床試験（RCT: Randomized Clinical Trial）であり、マスキングするのが通例である。RCTのようなパラレルデザインに対し、異なった介入に対する逐次のアウトカムを比較する逐次対照試験もあり、前後比較あるいは自己対照試験と呼ばれる。

支援機器の実証試験においてはマスキングが不可能である上に、2群に分けて比較して統計的に意味のあるデータを記録するのに十分な数の被験者を集めるのは困難であるし、それだけの数の試作支援機器を取りそろえるのはコスト的にも困難である。結局、前後比較試験が支援機器開発において最も信頼できる研究デザインと云える。

支援機器の実証試験においては、実質的には前後比較試験によらざるを得なくなる。実際には、対照のない介入実験もしばしば見受けられる。この場合、観測されたアウトカムからエンドポイントを導くのが容易ではなく、エビデンスとしての説得力は前後比較の場合に比べて弱い。

さらに、開発に先だって利用者のニーズを汲み取る段階や使用中の支援機器の問題点を掘り起こす段階があり、記述的観察研究から始まるのが通例である。

3. 支援機器の開発段階と実証試験の相

実証試験のプロトコルにおいてゴールが明確でない場合が見受けられるが、このことは、その実証試験において解決すべき問題点が十分に絞り切れていないためである。

医薬品や医療機器の開発においては治験に際してどの相の治験であるかを明確に定義する事が確立している。どの相の治験であるかによってゴールが定義される。

Independence Technology 社が iBOT の治験を FDA に申請するに当たっては、3段階のパイロット試験のあとで本格試験を行った⁽³⁾。このように支援機器の実証試験においても開発段階に応じたプロトコルを立案するのが適当であり、それによって実証試験のゴールをどう設定すべきかの指針とすることができる。

治療法や医薬品、医療機器の治験においては第 I 相から第 IV 相まで区分して行われるのが通例である。この考え方を支援機器に当てはめて表 1 の定義を提案したい。第 I

Table 1 Definition of phases in the development of medicine and of assistive products

相	治療法および医薬品・医療機器 ^(4,5)		支援機器
	主要点	目的	
第0相試験	2006年にFDAが提案。末期がん患者に少量の処方をして、薬物動態の観察を行うことによって新薬の開発を促進。		現在使用中の機器の使用状況の観察、支援すべき身体機能の特性測定などを目的とし、介入のない、または最小限の介入（現在使用中の機器の問題点を確認・探索する場合など）を伴う試験。
第I相試験	新治療法の人体・患者への最初の投与	さらに研究を進めることが適切かどうかを決定するための探索的研究。医薬品の開発の場合、健康成人に少量の投与により安全性や薬物動態を調べる。	試作した支援機器の健常者による安全性の確認。 [集会などにおける不特定多数の利用者による短時間の試用による適合・選好の確認]
第II相試験	患者における新治療の早期試験	用量反応の相関、毒性の発生率、潜在的要因の発見。医薬品の場合、比較的少数の患者を対象に有効性、副作用を検討。用法、用量を決定する。	想定する利用者の数名から10名程度の被験者によるパイロットテスト。有用性・適合・環境との整合性などの確認を目的とする。
第III相試験	新治療と標準治療との大規模比較試験	現在の標準治療に置き換えるべきかの最終評価。RCTの研究モデル。多数の被験者を対象とする	市場に出す前の最終確認。様々な条件下で20-40名程度の利用者による本格試験。有用性・適応・適合の実証を目的とする。
第IV相試験	新治療法を日常臨床で用いた場合のモニタリング	強固な副作用の発生率など、新治療の疾患に対するインパクトに関する追加情報のための市販後調査	市販後のフォローアップ。有害事象の解明、適応の拡大などを目的とする。

相から第IV相に関しては基本的には医薬品に関するものを踏襲しているが、第II相、第III相に関してサンプルサイズの目安を示した点が異なっている。これらの目安に対しては統計学的信頼性が十分ではないとの批判が寄せられるかも知れない。これは支援機器における被験者募集上の困難、試験のコストを考慮して決めた目安である。サンプルサイズが小さい点については生物学的妥当性に基づく考察によって補う必要がある。

第0相は2006年にFDAが抗がん剤の開発を促進するために提案したもの⁶⁾で、末期がん患者に少量の処方をして、薬物動態を観察しようとするものであった。

医薬品の開発においてはin vitro実験、動物実験などで薬効の可能性、安全性についての検討の後で人を対象とした研究段階に入るが、支援機器の開発においてはin vitro実験や動物実験を使うことはできず、人を対象とする観察研究から始めざるを得ない。そこで、開発を目的とした観察研究あるいは最小限の介入を伴う研究を第0相として位置づけることを提案したい。

このことによって開発の初期からきちんとしたプロトコルを策定することが推進できるし、この段階から倫理審査を受ける習慣をつける意義がある。なお、ほとんどの場合最小限の危険(minimal risk)しか伴わない研究と考えられ、迅速審査の手続きの可能性もある。

なお、第I相に集会などにおける不特定多数を対象とした試用について記載してあるが、開発段階によっては第II相試験として行われることもある。括弧をつけてあるのはそのためである。当然このような研究についても倫理審査が必要である。

4. 我が国における実証試験に関する検証

我が国における実証試験について、ゴールが明確でない、主観評価に頼りすぎているのではないかと懸念を表明したが、この印象を検証するための調査を行った。対象としたのは、リハ工学カンファレンスの2007年から2011年

までの5年分の予稿集に収録された論文である。

予稿集から支援機器開発に関連する論文137編を選定し、研究の性格として、第I相試験(Phase I test)、第II相試験(Phase II test)、第III相試験(Phase III test)、第IV相試験(Phase IV test)、開発症例報告(Case report)、開発事例報告(Development report)に分類して分析した。どの相に属するかは論文の内容から判断したもので、論文の著者の意図を表すものではない。なお、開発症例報告は利用者のニーズに基づいた開発で、被験者数も1名から数名までの小規模な開発研究である。開発事例研究はニーズから出発したもので、ハードウェアの開発に重点があり、被験者実験ではないものも含む。

4-1 研究デザイン

研究デザインとして、前後比較(Time series design)、対照のない介入(Intervention without control)、記述的横断研究(Descriptive cross sectional study)、症例研究(Case series)、開発研究(Development)に区分して分布を調べた。結果を表2Aに示す。

RCTが1件もなかったのはある意味では当然であるが、前後比較研究の研究デザインが総数で11件と少なく、対照のない介入の62件と比べてあまりに少ない事に驚く。

さらにながかりすることは、この表には示されていないが仮説検証型の本格的な実証試験の研究デザインが1件も含まれていなかったことである。これら2点は予稿集としての性格にもよっているかもしれないが、堅固な研究デザインに関する知識が普及していないでもあろう。実証試験のあり方について教育、普及の努力が不足していた点、反省要因である。

症例研究、開発研究などが多いのは臨床現場からの論文が多い事から当然予測されることであるが、同時に萌芽的な開発が活発であることも意味している。これらについては、現段階にとどまることなく実用化に向けての今後の発展に期待したい。

4-2 評価関数

主観評価が多いという印象を検証するために対象とした論文に採用されている評価関数に着目した。集計した結果を表 2.B に示す。評価の客観性については、客観指標あり (Objective outcome), 被験者主観評価 (Subjective evaluation by subjects), 研究者主観評価 (Subjective evaluation by researchers), 中間ユーザ主観評価 (Subjective evaluation by professionals), アウトカムなし (Outcome not defined), 評価無し (No evaluation) に分類した。

客観指標有りはアウトカム変数に客観指標が含まれているものすべてを区分したので、最終判定が客観指標に基づいている訳ではない。何らかの客観指標を含むものが総数の 4 分の 1 しかなく、主観評価が多いという印象を裏付けている。

主観評価、特に被験者および研究者自身による主観評価はバイアスや交絡因子を伴いやすく、可能な限り避けるべきであり、我が国における支援機器の科学的な開発のためには更に改善が求められる。

4-3 倫理審査

倫理審査についての記載状況を表 2.C にまとめた。倫理審査の状況は審査済み (Approved), 説明と同意 (Informed consent), 要審査 (Not described), minimal risk (May be minimal risk), 診療 (Practice), 行政目的 (Administrative purpose), 試作報告 (Prototype development) に区分した。

説明と同意は説明の上で同意を得たとの記載があるもの、minimal risk は最小限の危険しか含まないので迅速審査が適用されると判断したもので、要審査の内数である。審査不要のうち、診療は Belmont Report の原則によって倫理審査の対象とはならない。試作報告は主としてハードウェアの開発を中心とした報告であり人を対象とした研究ではないので倫理審査の対象とはならない。行政目的の場合は倫理審査の対象とはならないと国際的にも考えられている。

審査済みとの記載があるのは 3 件に過ぎずあまりに少な

いと印象を受けるが、この予稿集においては倫理審査に関する記述を要求してはならないので記載がなくとも審査を受けていないことを表すわけではない。しかし、説明の上同意を得たとの記載を含めても研究倫理に関する配慮が不足しているとの印象は変わらない。

倫理審査は科学面と倫理面の両方から行われるので、本報告で検討したプロトコルの科学性についても倫理審査で検討されるべきである。これについても強調しておきたい。

5. おわりに

我が国における支援機器実証試験の科学性を増進させるために、前後比較の研究デザインの重要性、実証試験の相に基づいたプロトコルの提案を行うとともに、その実情に関する調査を行った。科学性が不足しているとの印象を裏付ける結果となった。よりいっそうの努力が求められる。

参考文献 (URL のアクセスは 2012 年 9 月 10 日)

- (1) John I. Gallin, ed. (井村裕夫, 竹内正弘, 藤原康弘, 渡辺亨監訳), NIH臨床研究の基本と実際, 丸善, 2004.
- (2) Stephen B. Hulley他, (木原雅子, 木原正博訳) 医学的研究のデザイン第3版, メディカル・サイエンス・インターナショナル, 2009.
- (3) 山内繁, iBOTの臨床試験をめぐって, 日本生活支援工学会誌, vol 11, no 2, pp3-9, 2011.
- (4) John I. Gallin, ed. (井村裕夫, 竹内正弘, 藤原康弘, 渡辺亨監訳), NIH臨床研究の基本と実際, 丸善, p. 206, 2004.
- (5) E. G. DeRenzo and J. M. Moss, "Writing Clinical Research Protocols: Ethical Considerations", p.57-8, Elsevier, 2006.
- (6) FDA, "Guidance for Industry, Investigators, and Reviewers, Exploratory IND Studies", 2006.
<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm078933.pdf>.

Table 2 Summary of Protocols of papers on R&D of Assistive Products in Proceedings of RESJA Conference, 2007-2011

A. Research Design							
	Phase I test	Phase II test	Phase III test	Phase IV test	Case report	Development report	Total
Time series design	2	4	1	1	1	2	11
Intervention without control	16	17	5	4		20	62
Descriptive cross sectional study				2		2	4
Case series					41		41
Development					3	13	16
B. Evaluation Functions							
Objective outcome	10	8	1	2	5	11	37
Subjective evaluation by subjects	5	8	3	3	7	6	32
Subjective evaluation by researchers		2	1		20	8	31
Subjective evaluation by professionals	1	1					2
Outcome not defined	1	1			4	3	9
No evaluation	1	1	1	2	9	9	23
C. IRB Review							
IRB review mandatory							
Approved	2					1	3
Informed consent		6		1		3	10
Not described (May be minimal risk)	16 (2)	15 (2)	6	5		25	63 (4)
IRB review not mandatory							
Practice					45	4	49
Administrative purpose				1			1
Prototype development						4	4
total	18	21	6	7	45	37	134