

NGO インパクト評価 10 ステップ

— 質問票とエクセルの操作手順付きですぐ使える小規模事業インパクト評価ガイドライン —

Impact Evaluation Guideline for NGO Project



第 1.3 版

2017 年 12 月改訂

執筆者代表：佐々木亮（評価学博士）

国際開発センター（IDCJ）評価部

なぜインパクト評価が必要か？

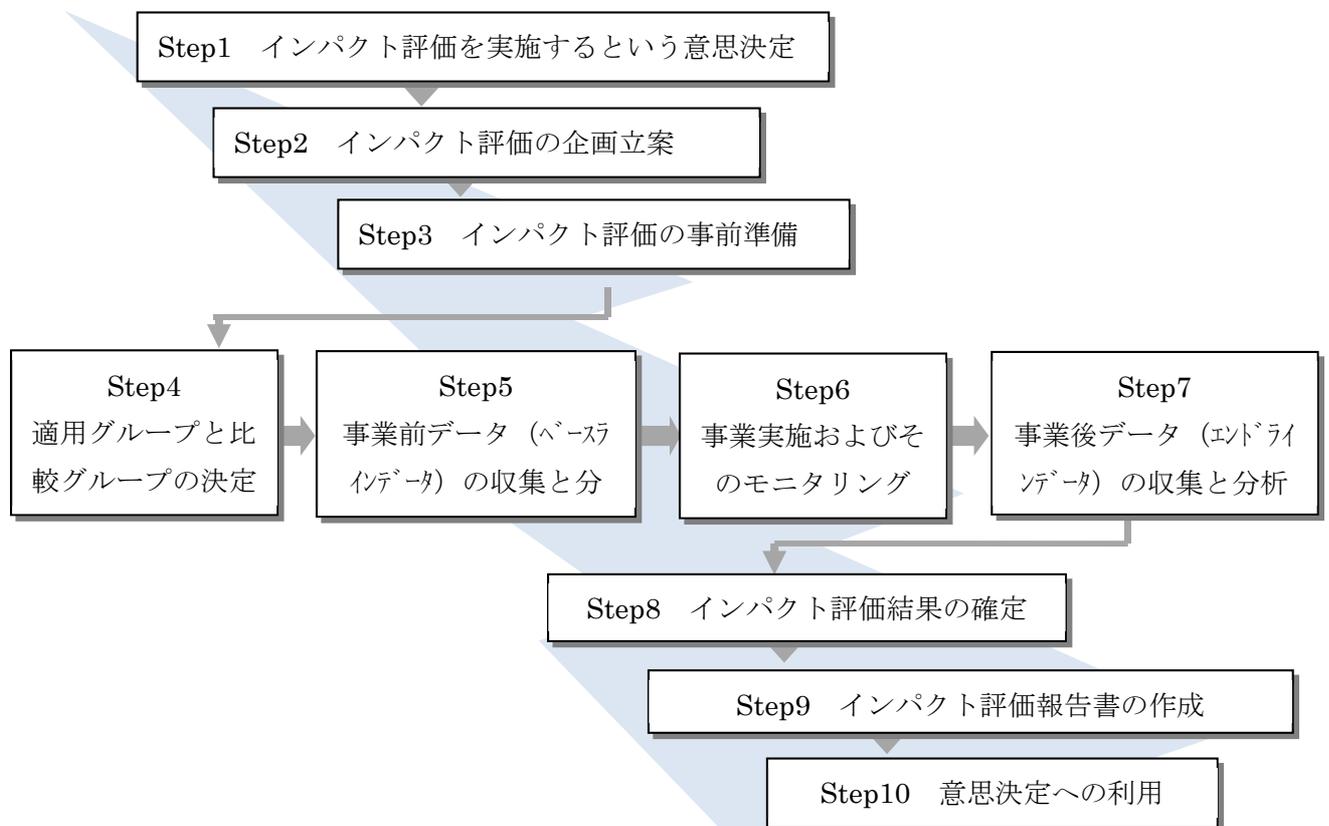
近年、NGO/NPO の間で「インパクト評価」が注目されるようになりました。それは次のような必要性に対応することを迫られているからだとは見えています。

- ・新しい事業を思いついたが効果があるのかどうかを確かめる必要がある。
- ・資金支援者に、支援を受けて実施している事業は効果があると示す必要がある。
- ・個人会員の数および寄付額を増加させるために事業効果を分かりやすく示す必要がある。

これらの必要性に応える「インパクト評価」は、過去の学術研究や援助機関の実務経験に裏打ちされた標準的な手続きが確立されています。その標準的な手続きをなるべく分かりやすくとりまとめたのがこのガイドラインです。このガイドラインが広く普及することによって、NGO/NPO の間で適切な形でインパクト評価が普及し、そして組織の意思決定と説明責任の向上が図られることを期待しています。

なお、このガイドラインは、国際開発センターが主催した「NGO 事業評価勉強会」(2012年)の議論を通じて作成されました。末筆になりましたが、同勉強会にご参加いただき事例発表およびガイドライン案へのコメントをお寄せいただいた NGO 有志の方々にこの場を借りて深く御礼申し上げます。

執筆者を代表して
佐々木 亮



NGO インパクト評価 10 ステップ

—質問票とエクセルの操作手順付きですぐ使える小規模事業インパクト評価ガイドライン—

目次

はじめに：インパクト評価が必要になる状況とは？	1
ステップ1：インパクト評価を実施するという意思決定	3
ステップ2：インパクト評価の企画立案	7
ステップ3：インパクト評価の事前準備	9
ステップ4：適用グループと比較グループの決定	11
ステップ5：事業前データ（ベースラインデータ）の収集と分析	17
ステップ6：事業実施およびそのモニタリング	21
ステップ7：事業後データ（エンドラインデータ）の収集と分析	23
ステップ8：インパクト評価結果の確定	26
ステップ9：インパクト評価報告書の作成	27
ステップ10：意思決定への利用	28
別添1：質問票（雛形）	29
別添2：エクセル入力フォーム（雛形）	30
別添3：エクセル操作法（等分散の検定、2群の有意差検定）	31
別添4：エフェクトサイズの計算	38
著者等の紹介	39

はじめに：インパクト評価が必要になる状況とは？

国際協力 NGO や特定非営利活動法人(NPO)においてインパクト評価が必要となる状況は多々あるが、代表的な状況は次のような場合である。

- ・新しい事業を思いついたが効果があるのかどうかを確かめる必要がある。
- ・資金支援者に、支援を受けて実施している事業は効果があると示す必要がある。
- ・減少傾向が続く個人会員の数および寄付額を増加させるために事業効果を分かりやすく示す必要がある。

従来は、新しい事業のアイデアを思いついた場合に、その効果を確認せずに理事長や理事会の承認のみで実施するケースが多かった。その結果、当該事業が 2~3 サイクルが終わる頃に、事業の効果の有無が問題となることがしばしばあった。こうした状況を避けるために、事業の本格的な実施段階の前に、時間と場所を区切って、効果が確かにあるのか検証してみることの意義が広く認められるようになってきた（「パイロット事業」と呼ぶこともある）。

もうひとつの状況は、NGO であれ NPO であれ、公的な資金支援者（省庁、公益財団）や CSR 企業から支援事業の効果を明示してほしいと依頼される場合である。以前は、資金支援者が、資金支援の総額を減額したり支援事業の数を絞り込んだりする場合に、事業効果の明示が求められることが多かった。しかし最近では、あらかじめ全ての事業に事業効果を明示することを義務付けることが一般的となりつつある。こうした「事業効果の明示義務」に応えるためにインパクト評価の実施が必須となってきている。

さらにもうひとつの状況は、個人会員の数および寄付額が継続的に減少している現状を改善する必要がある場合である。NGO も NPO も設立当初は、熱心な個人会員やサポーターからの会費収入や寄付金収入で安定した活動資金を確保できる場合があるが、年数が経つと、目新しさが失われるとともに、新興の NGO/NPO に押されて、会員数および寄付額が次第に減少していき、そしてついに活動資金に支障をきたすような状況に陥ることもある。こうした状況から脱するために、貴重な活動資金を使って行った事業の事業効果を明確にわかりやすく、既存の、そして潜在的な個人会員やサポーターに示さねばならなくなる。こうした状況ではインパクト評価の実施が必須となる。

なお、以下に「はじめに」にインパクト評価実施における全体の流れ図をしめしたので参照されたい。

そもそもインパクト評価とは何か？

そもそも「評価」の定義としてもっとも適切なのは「評価はものごとの価値を明らかにすることである」という定義である（(Scriven, M. 1991, p.1. “Evaluation is determination of merit, worth and significance of things.”。この定義は、単なる事実特定に留まらずに、特定された事実に基づいて何らかの価値判断を下してはじめて評価足りえるということを端的に示している。

調査（Research）が事実特定に主眼を置いているのに対して、評価(Evaluation)は特定された事実に基づいて何らかの価値判断を下すことに主眼を置いている。

$$\begin{array}{ccccc} \boxed{\text{評 価}} & = & \boxed{\text{事実特定}} & + & \boxed{\text{価値判断}} \\ \text{(Evaluation)} & & \text{(Factual identification)} & & \text{(Value determination)} \end{array}$$

インパクト評価もこれに沿う。したがって、第一に、効果（インパクト）の有無と程度という事実関係を客観的に明らかにして、第二に、あらかじめ設定された価値基準に基づいて、「たいへん満足できる」「満足できる」「満足できない」などの価値を表現する言葉で明確に価値判断をする必要がある。

ステップ1：インパクト評価を実施するという意思決定

インパクト評価が必要となった状況を踏まえて、最初に「インパクト評価を実施する」という組織としての意思決定がなされる必要がある。

1.1 新しい事業を思いついた場合：パイロット事業として実施する

NGO/NPO 内部で職員から新しい事業のアイデアに関する提案があった場合、いきなり本格実施をせずに、まずは場所を限定して本当に効果があるかどうかを確かめることが勧められる。まずは、パイロット事業として実施するということが理事会で提案されて承認される必要がある。承認されれば、次年度の活動計画案および予算案にその実施に必要な人員と資金が盛り込まれ、その活動計画案および予算案が承認されて、晴れてパイロット事業として実施できることになる。

1.2 資金支援者や個人会員に対する説明責任を果たす場合：実施中の事業を評価する

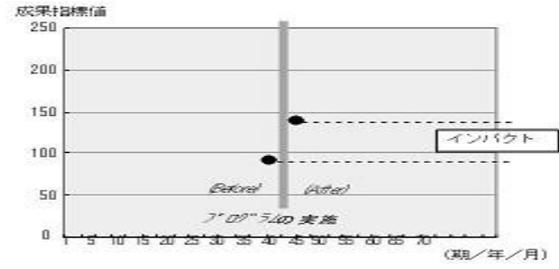
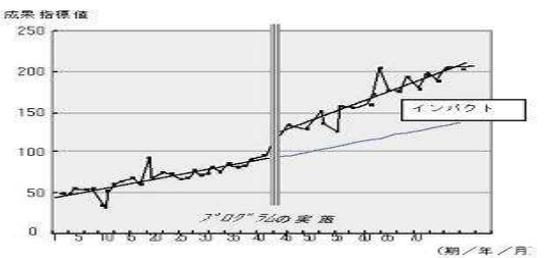
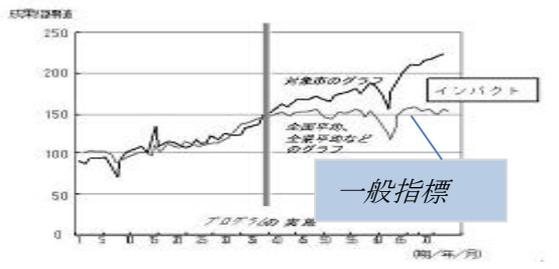
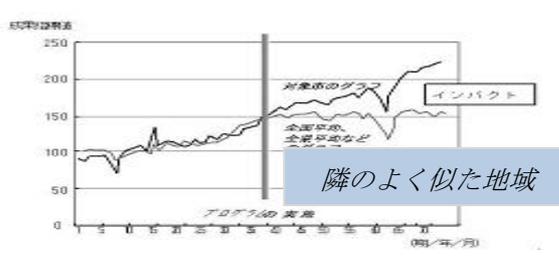
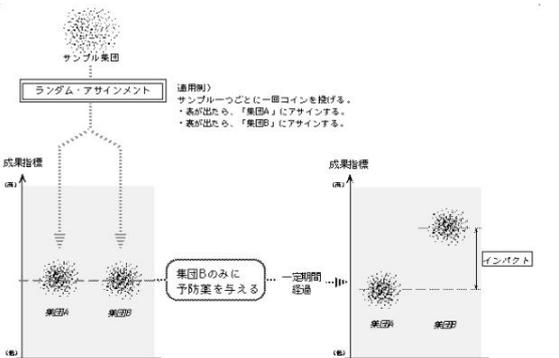
資金支援者（省庁、公益法人、CSR 企業）や個人会員（いわゆるサポーターを含む）に対して事業効果を明示するために、インパクト評価を実施するという提案が理事会でなされて承認される。次年度の活動計画案および予算案に、実施地域と比較地域の双方からのデータ収集のための人員と資金が盛り込まれ、その活動計画案および予算案が承認されて、晴れてインパクト評価が実施されることになる。

なお、公的支援者や CSR 企業からは、すでに事業の本格実施の承認を得ている場合が多いであろうから、上記のようにパイロット事業としての実施を考えることは適切ではない。その場合には、何らかの方法により比較地域を定めて、事業を実施する地域と、実施しない地域との間の各種指標値の差を測定することによりインパクトを推定する。

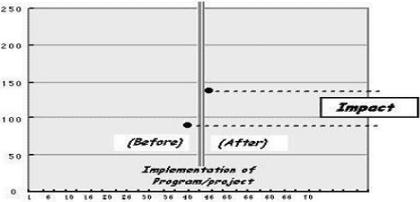
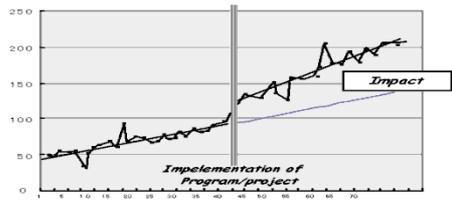
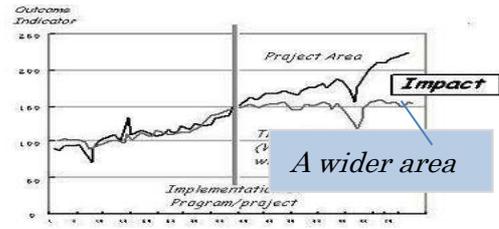
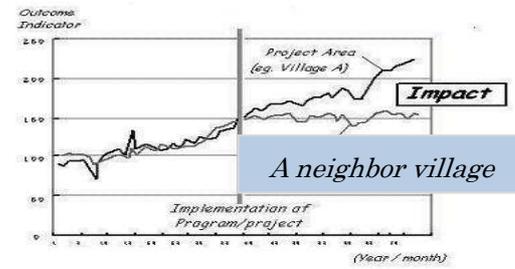
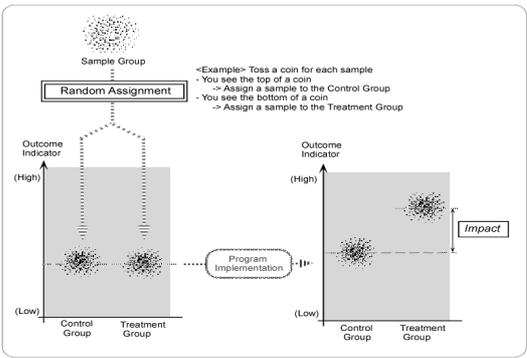
現状では、地域全体のデータによる単純な事前・事後比較デザインが頻繁に用いられているが、客観的に効果を示すためには非常に脆弱なデザインであるため、率直に言ってあまりお勧めできない。

参考までに、以下に、インパクト評価の基本的デザインを示す。たいへん簡便かつローコストで実施できる「事前-事後デザイン」から、厳格だけれども高額な費用と多大な時間が掛かることが多いランダム実験デザイン（Randomized Controlled Trial (RCT)）まで、代表的な 5 種類を和文、英文で簡単に解説している。

インパクト評価の基本デザイン

名称と概念図	説明	
<p>1. 事前・事後比較デザイン(Before-After Design)</p> 	<p>シンプルに、事前、事後の指標値を比較し、差があれば因果関係があったと推定する。簡便なので広く用いられている。ただし、事前・事後の間に発生した外部要因による影響値をまったく取り除けないので、因果関係の推定の信頼性は低い。 [検定テスト:対応のある(一対の)t検定]</p>	<p>単純 & 安価</p> <p style="text-align: center;">↑</p>
<p>2. 時系列デザイン (Interrupted Time-Series)</p> 	<p>施策介入前の長期的トレンドを導き出し、施策介入後にトレンドが変わっていれば、因果関係の存在を推定する。ただし、長期的トレンド以外の外部要因による影響値を取り除けないので、信頼性はそれほど高くない。 [検定テスト: 回帰分析]</p>	<p style="text-align: center;">↑</p>
<p>3. 一般指標デザイン (Generic Control)</p> 	<p>全国平均値、全県平均値などの一般指標値を比較に用いる。外部要因による影響値をある程度除去して考えることができるので(なぜなら対象地域が受けた影響とある程度同じ影響を一般指標値も受けているはずだから)、因果関係の存在の特定に関してある程度の信頼性を確保できる。わりと簡単に用いることができる。 [検定テスト: 目視による判断]</p>	<p style="text-align: center;">↑</p>
<p>4. マッチングデザイン (Matched control)</p> 	<p>可能な限り近似のグループを選定して比較に用いる。外部要因による影響はどちらのグループも同程度に受けると考えられるので、因果関係の存在の特定のために高い信頼性を確保できる。地域全体の特徴でマッチングする「地域マッチング」と、個人を一人一人マッチングする「個人マッチング」がある。 [検定テスト: 独立のt検定]</p>	<p style="text-align: center;">↑</p>
<p>5. ランダム実験デザイン (Randomized controlled trial (RCT))</p> 	<p>施策の実施前に、政策適用を無作為割付(ランダム・アサインメント)により、実施グループと比較グループに分ける。成果指標(Outcome indicators)に現れた違いは、途中の唯一の違いである「介入を適用されたか否か」によって引き起こされたと純粋に判断することができる。因果関係の存在の特定について高い信頼性を誇る。ただし実際の適用は難しい。 [検定テスト: 独立のt検定]</p>	<p>厳格 but 高価</p> <p style="text-align: center;">↓</p>

Basic designs of impact evaluation by graphical presentation

Name & graphical concept	Explanation	
<p>2. Simple before-after design</p> 	<p>You designate <u>one point in time <i>before</i> the program/project was conducted</u> and <u>another point <i>after</i> completion</u>. Then you estimate the difference as impact. [Test: Dependent (Paired) t-test]</p>	<p>Simple and less costly</p>
<p>2. Interrupted Time-series design</p> 	<p><u>The long-term data before the program is conducted</u> as well as <u>the long-term data after the program is conducted</u> are collected. Then, a trend line for each data is drawn and the discrepancy between the two trend lines is estimated as impact of the program. [Test: Regression analysis]</p>	
<p>3. Generic control design</p> 	<p>In this design, you use <u>the average of a wider area as the comparison</u> ---- the whole country, whole region or whole prefecture to which the project area belongs. Then you estimate the difference as impact. [Test: Your eye bowling judgment]</p>	
<p>4. Matched control design</p> 	<p>In short, <u>compare your village with a neighbor village</u>. The difference is regarded as impact. The point is how much the two villages are identical (or comparable) except whether the intervention is applied or not. [Test: Independent t-test]</p>	
<p>5. Randomized controlled trial (RCT)</p> 	<p>By using random assignment, for example by tossing a coin, two groups are created. They are <u>totally statistically identical (or comparable)</u>, and they are expected to <u>be affected by external events to exactly the same degree</u>. Thus, it can be judged that any difference resulted on any indicators were caused by the only difference between them through the process, which is whether the intervention was applied or not. [Test: Independent t-test]</p>	<p>Rigorous but expensive</p>

ケーススタディ A: マイクロファイナンス事業（ランダム化比較試験（RCT）の例）

横浜に本部を置く国際 NGO「国境なき隣人の会」（仮称）は 1970 年代の設立以来、ネパールの農村女性たちの自立を支援してきた。主な活動は、ミシンの寄贈である。同会の活動費は設立以来のメンバーである会員からの寄付金で成り立っているが、近年、会員数は頭打ちであり、新規の若い会員の獲得が長年の課題ともなっている。

今般、新しくこの NGO の事務局長に就任した阿部さん（仮名）は、ネパールにおいて、農村女性向けのマイクロファイナンス事業を本格的に開始したいと考えている。近い将来、理事会でマイクロファイナンス事業の開始の承認を得るために、とりあえずネパール南部の平野部のシンズリ県の農村地帯で期間限定（1年半）のパイロットプロジェクトとして実施することになった。



ケーススタディ B: 地場野菜の共同出荷事業（地域マッチングの例）

山形に本部を置く国際 NGO「かるがも！」（仮称）は、2005 年に山形の大学生達が設立した比較的若い国際 NGO である。カンボジアの農村でコミュニティ開発を支援しており、具体的には農産物の共同出荷を支援している。週一回小型トラックで、近くの地方都市の市場へ地場野菜を出荷する。これにより農民世帯の収入向上を実現することを目指している。

今までは現地の NGO と協力して細々と活動を続けてきた「かるがも！」であるが、今までの活動に手ごたえを感じ、今後、現地に事務所を借りて専従職員 1～2 名を派遣して、日本から定期的に大学生インターンを派遣して支援活動を拡大したいと考えている。そのためには、日本国内での寄付金の獲得や、公益財団や公的機関からの補助金獲得が必須である。そのため、来年度新たに共同出荷事業（まずは 1 年間の実施）を開始する村落に関して簡便なインパクト評価を実施しようという話が持ち上がった。インパクト評価を実施して、地場野菜の共同出荷事業が本当に効果があることを証明し、その評価結果を広報に使ったり、補助金申請書に記載して説得材料とすることを考えている。



ステップ2：インパクト評価の企画立案

インパクト評価を実施するという意思決定がなされたら、さっそくインパクト評価の企画立案にとりかかる。決めなければならないことは次の諸点である。

- ・インパクト評価の目的
- ・評価しようとする事業の対象地域・対象人口
- ・インパクト評価の基本的デザイン（マッチングデザインか RCT か、など）
- ・実施者（内部評価か第三者評価か）
- ・予算と人員
- ・スケジュールと実施手順

以上の点を企画書に記載する。具体的な記載内容は以下の通りである。

2.1 インパクト評価の目的

「〇〇事業の実施により、〇〇の状況の改善度合いを評価し、次年度以降の事業継続の可否を判断する」、「〇〇事業の実施により、〇〇の状況の改善度合いを評価し、資金支援者／個人会員に対する説明責任を果たす」と記載する。

2.2 評価しようとする事業の対象地域・対象人口

インパクト評価の対象地域あるいは対象人口を明確に記載する。パイロット事業としての実施であればその対象地域を明確に記載する。すでに本格実施している事業であれば、あとで比較地域（非実施地域）が特定できるように、その事業の対象地域を明確に記載する。

2.3 インパクト評価の基本的デザイン（マッチングデザインか RCT か、など）

もっとも厳格なランダム化比較試験（RCT：Randomized Controlled Trial）、マッチングデザイン（Matching Design）、一般指標デザイン（Generic Control Design）などからどれを用いるかを明記する。ただし、もっとも広く用いられている事前・事後比較デザインは、客観性が低く勧められない。

2.4 実施者（内部評価か第三者評価か）

実施者を決定する。大きく分けると内部評価と外部評価（第三者評価）がある。NGO/NPOの職員が実施する評価は内部評価であり、独立の第三者に委託して行う第三者評価よりも、恣意的になりやすいと言われる。ただし内部評価の方が実施コストは格段に低い。外部専門家と内部者による「合同評価」もある。

2.5 予算と人員

旅費宿泊費・現地協力者謝金・評価者謝金（外部評価者に委託する場合）を見積もる。人員は、評価者1名（あるいはチームで2名）、現地協力者（現地で訪問調査をする現地人数

名)を見込むのが普通である。

2.6 スケジュールと実施手順

内部評価であれ外部評価であれ、実施期間は 2 ヶ月あるいは 3 ヶ月程度を見込むべきである。現地調査期間は 2 週間程度で終了させるのが現実的である。

ケーススタディ A: マイクロファイナンス事業 (RCT)

国際 NGO 「国境なき隣人の会」の阿部さんは、次のとおりの企画書をまとめた。

- ・インパクト評価の目的: 「マイクロファイナンス事業の実施により、農村世帯における女性の地位向上および社会経済状況の改善度合いを評価し、次年度以降の本格実施の可否を判断する」
- ・評価しようとする事業の対象地域・対象人口: ネパール国シズリ県北東地方の 3 つの村
- ・インパクト評価の基本的デザイン: 最も厳格な RCT (ランダム化比較試験) を適用
- ・実施者: 外部者と内部者の「合同評価」で実施する。阿部事務局長と XX 大学グローバル研究所の佐藤准教授 (仮名) とが共同の実施者。
- ・スケジュールと実施手順: 現地調査は 2 回実施する。 2010 年 8-9 月 (事前段階の作業期間) と 2012 年 2-3 月 (事後段階の作業期間) (佐藤准教授の大学の長期休暇を利用する)
- ・予算と人員: 2010 年と、2012 年の 2 回分の旅費宿泊費・現地協力者謝金・評価者謝金を含めて約 300 万円。

ケーススタディ B: 地場野菜の共同出荷事業 (地域マッチング)

国際 NGO 「かるがも！」(仮称) は、次のとおりの企画書をまとめた。

- ・インパクト評価の目的: 「地場野菜の共同出荷事業の実施により、農民の収入向上と生活状況の改善度合いを評価し、資金支援者/個人会員に対する説明責任を果たす」
- ・評価しようとする事業の対象地域・対象人口: カンボジア国チャオプラヤ県 XXX 村
- ・インパクト評価の基本的デザイン: マッチングデザイン (地域マッチング) を適用
- ・実施者: 内部評価で実施する。「かるがも！」の鴨下理事 (非常勤) (仮名) が実施する。
- ・スケジュールと実施手順: 現地調査は 2 回実施する。 2010 年 8-9 月 (事前段階の作業期間) と 2011 年 8-9 月 (事後段階の作業期間) (それぞれ 3 週間ずつ)
- ・予算と人員: 2010 年と 2011 年の 2 回分の旅費宿泊費・現地協力者謝金・評価者謝金を含めて約 100 万円。コストが安いのは、内部の非常勤の理事に依頼する内部評価なので評価者謝金が非常に低額で済むからである。

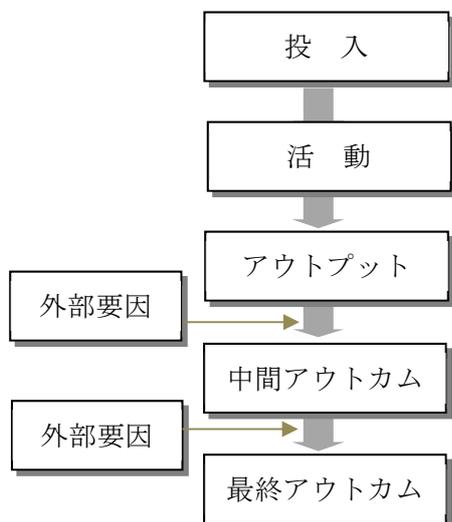
ステップ3：インパクト評価の事前準備

事前準備として、現地調査前に以下の項目を実施する。

- ・ロジックモデルの作成
- ・アウトプット指標、アウトカム指標の設定
- ・収集すべきデータと収集方法の決定
- ・質問票（クエスチョネア）の作成
- ・評価枠組（評価項目×評価基準で構成）

3.1 ロジックモデルの作成

事業の介入理論を表現するフローチャートを作成する。



3.2 アウトプット指標、アウトカム指標の設定

ロジックモデルで特定されたアウトプット、アウトカムの実現度合いを適切に示す指標を設定する。アウトプット、中間アウトカム、最終アウトカムのそれぞれのレベルの指標は1つである必要はなく、必要に応じて3つ程度まで設定しても構わない。以下が一例である。

【アウトプット指標】

*受益者であるかどうかの再確認のためこの質問をする。
「〇〇生産研修を修了した」⇒指標：Yes=1, No=0

【中間アウトカム指標】

「〇〇の家内生産を始める」⇒指標：生産量（kg/ヶ月）

【最終アウトカム指標】

「収入が増加する」⇒指標：収入額（ドル）

「自尊心が高まる」⇒指標：自尊心の度合（5段階）

3.3 収集すべきデータと収集方法の決定

主な方法は、質問票を用いた個別インタビュー、フォーカスグループインタビュー、関係者インタビュー、既存統計資料の収集などである。適宜組み合わせる。

3.4 質問票（クエスチョネア）の作成

アウトプット、アウトカム、回答者の属性に絞った最小限の質問票にすべきである。聞いておいたら役に立つかも知れないという質問はなるべく入れない。

3.5 評価枠組（評価項目×評価基準で構成）

インパクト評価の場合の「評価項目」はインパクト（あるいは有効性）のみであり、評価基準は、「A:大きな効果があった」「B:効果があった」「C:ある程度効果があった」「D:効果があったとは言えない」の4水準とするのが一般的である。

ケーススタディ A: マイクロファイナンス事業 (RCT)

国際 NGO「国境なき隣人の会」(仮称) の阿部さんは、次のとおりに準備を進めた。

<ul style="list-style-type: none"> ロジックモデルの作成 以下のとおりの想定とした。 事業名：マイクロファイナンス事業 		<ul style="list-style-type: none"> アウトプット指標、アウトカム指標の設定 	
投入	<ul style="list-style-type: none"> 人員(NGO 側、地元側) 資金 (融資の原資) 事務所スペース 	アウトプット指標	<ul style="list-style-type: none"> 融資を受けた人数 (人) 融資総額(ルピー)
活動	<ul style="list-style-type: none"> 農村女性向けの研修 融資申請受付・審査事務 貸出実行・返済受付事務 	短期アウトカム指標	<ul style="list-style-type: none"> 新規事業の件数 (件) (業種・規模も集計する)
アウトプット	<ul style="list-style-type: none"> 農村女性が融資を受ける (=資金を得る) 	中期アウトカム指標	<ul style="list-style-type: none"> 世帯の収入額 (ルピー)
短期アウトカム	<ul style="list-style-type: none"> 農村女性によって新規事業が実施される 	長期アウトカム指標	<ul style="list-style-type: none"> 「女性が世帯における主要な意思決定者か？」(5段階) 世帯全体の教育指標 (最終学歴、識字率)、保健指標
中期アウトカム	<ul style="list-style-type: none"> 農村世帯の収入が増加する 	<ul style="list-style-type: none"> 収集すべきデータと収集方法の決定 質問票調査 (農村女性への個別インタビュー) 質問票の作成 1 時間以内に終わる質問票を作成した。 評価枠組 エフェクトサイズを用いてインパクトの大・中・小を判断する仮の評価枠組を設定した。 	
長期アウトカム	<ul style="list-style-type: none"> 農村女性の地位が向上する 世帯全体の教育・保健状況が改善する 		

ケーススタディ B: 地場野菜の共同出荷事業 (地域マッチング)

国際 NGO「かるがも！」(仮称) の鴨下さんは、次のとおりに準備を進めた。

<ul style="list-style-type: none"> ロジックモデルの作成 以下のとおりの想定とした。 事業名：地場野菜の共同出荷事業 		<ul style="list-style-type: none"> アウトプット指標、アウトカム指標の設定 	
投入	<ul style="list-style-type: none"> 人員(NGO 側、地元側) 資金 事務所スペース 	アウトプット指標	<ul style="list-style-type: none"> 共同出荷の回数 共同出荷の品目数・量
活動	<ul style="list-style-type: none"> 農民向けの研修 小型トラック借受 共同出荷の実行 	短期アウトカム指標	<ul style="list-style-type: none"> 売上高 (品目別) 収入額 (ドル)
アウトプット	<ul style="list-style-type: none"> 共同出荷が実施される 	長期アウトカム指標	<ul style="list-style-type: none"> 耐久消費財 (テレビ、バイク) の有無 生活の満足度 (5段階)
短期アウトカム	<ul style="list-style-type: none"> 地場野菜の売上増加 	<ul style="list-style-type: none"> 収集すべきデータと収集方法の決定 質問票調査 (農民への個別インタビュー) 質問票の作成 1 時間以内に終わる質問票を作成した。 評価枠組 エフェクトサイズを用いてインパクトの大・中・小を判断する仮の評価枠組を設定した。 	
長期アウトカム	<ul style="list-style-type: none"> 農民の収入増加 生活の改善 		

ステップ4：適用グループと比較グループの決定

次に事業の適用グループと比較グループの決定を行う。そのためには以下の決定が必要である。

- ・ サンプルサイズの決定
- ・ 適用・比較グループの決定方法1：ランダムアサインメント（無作為割付）
- ・ 適用・比較グループの決定方法2：マッチングデザイン

4.1 サンプルサイズの決定

サンプルサイズを決定する公式はいくつかあるが、実際には予算の制約によって決定せざるを得ない。もちろん多ければ多いほどいいが、最低のサンプル数は1グループあたり30人あるいは25人であると言われる。以下が目安であり、あとは予算と相談して決定する。なお、それは1グループあたりのサンプル数であり、2群の場合にはその倍となる。

1グループのサンプル数	説明
30(あるいは25)	効果(インパクト)が大きいことが予想される場合の最低限のサンプル数 ¹ 。
60(あるいは50)	効果(インパクト)が中程度であることが予想される場合の最低限のサンプル数 ² 。
100	公式を使わない場合における望ましい最低限のサンプル数。なお、100取っていれば、あとで小グループに分割する際にも最低限の基準を満たしている可能性が高い。年齢幅などの特徴で3グループに分割すれば1グループあたり約30人、4グループに分割すれば1グループあたり約25人になっている。

4.2 適用・比較グループの決定方法1：ランダムアサインメント（無作為割付）(RCT³)

単純にコインの裏と表で2群に分ける方法である。このランダムアサインメント（無作為割付）により、あらゆる指標の平均値がほぼ同一である2つのグループを形成することが可能であり、この方法に勝る方法はないと言われている。インパクト評価における黄金標準（Gold Standard）と言われることもある。

実際の手続きには、エクセルの「分析ツール」の「サンプリング」(Sampling)を使う。次の例は、あらかじめ用意した60人を、ランダムアサインメントによって、30人と30人の

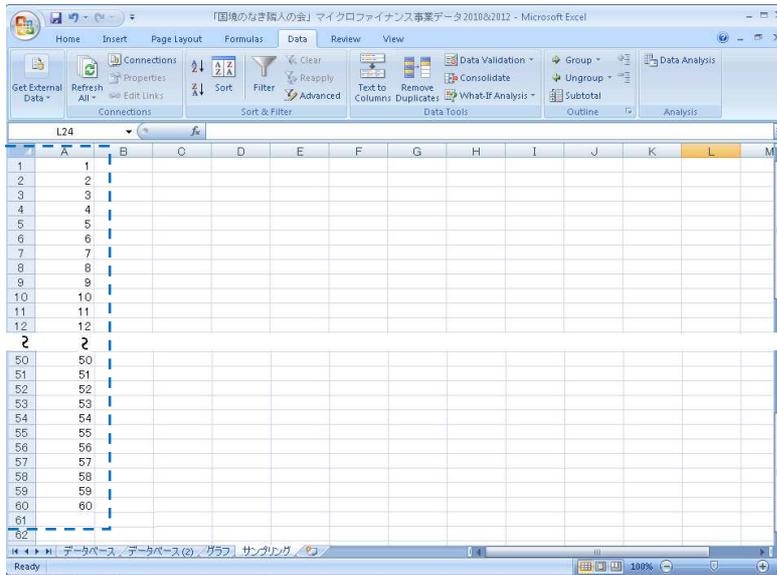
¹ パワー分析でサンプル数を計算する際に、 d (効果サイズ)=0.8(“大”), $\alpha=0.05$, Power $(1-\beta)=0.8$ である場合の計算結果は26となり、この値と近似している。

² パワー分析でサンプル数を計算する際に、 d (効果サイズ)=0.5(“中”), $\alpha=0.05$, Power $(1-\beta)=0.8$ である場合の計算結果は64となり、この値と近似している。

³ Randomized Controlled Trial の略。ランダム実験モデル（龍・佐々木(2004)）とも訳す。

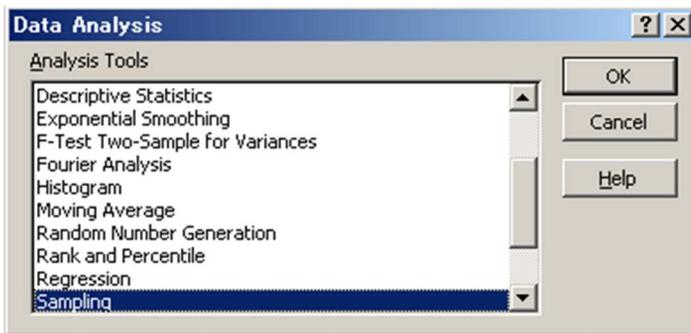
二つのグループに分ける際の例である。

① 最初に、ID 番号として、1 から 60 まで番号を振ったデータを作成する。

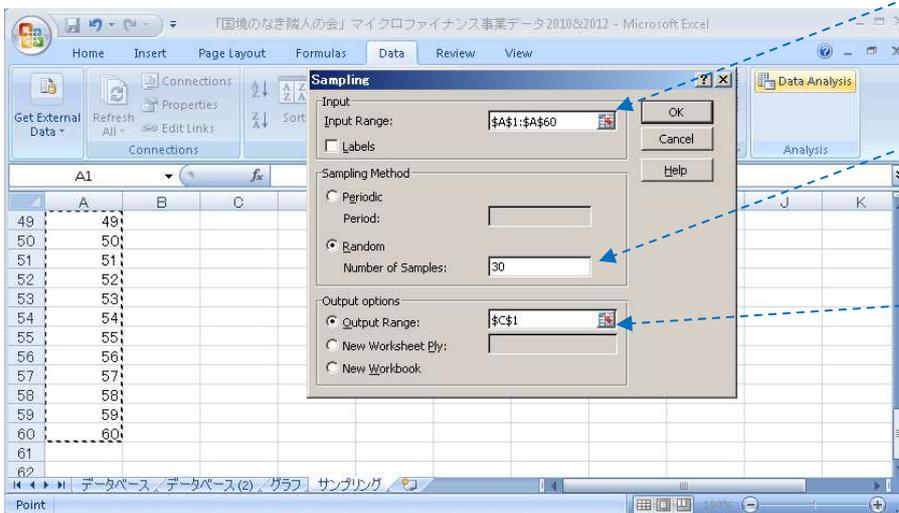


② 「データ」 → 「データ分析」 → 「サンプリング」 (Sampling) → 「OK」と選択する。

(*エクセルで各種の統計分析を行うためには、「分析ツール」というソフトをインストールする必要があります)



③ 現れたボックスに必要な情報を入力する。



入力した 1 から 60 までのセルを選択する。

30 個の ID 番号をランダムに得たいので、ここに 30 と入力する。

今見ているシートの余白に結果を生成させたいので、空いているセルを選択する。(この場合、C1 を選択している。)

- ③ 「OK」を押すと、30 個の ID 番号が得られる。この 30 人が実施グループに割り当てられる。残りの 30 人が比較（非実施）グループに割り当てられる。

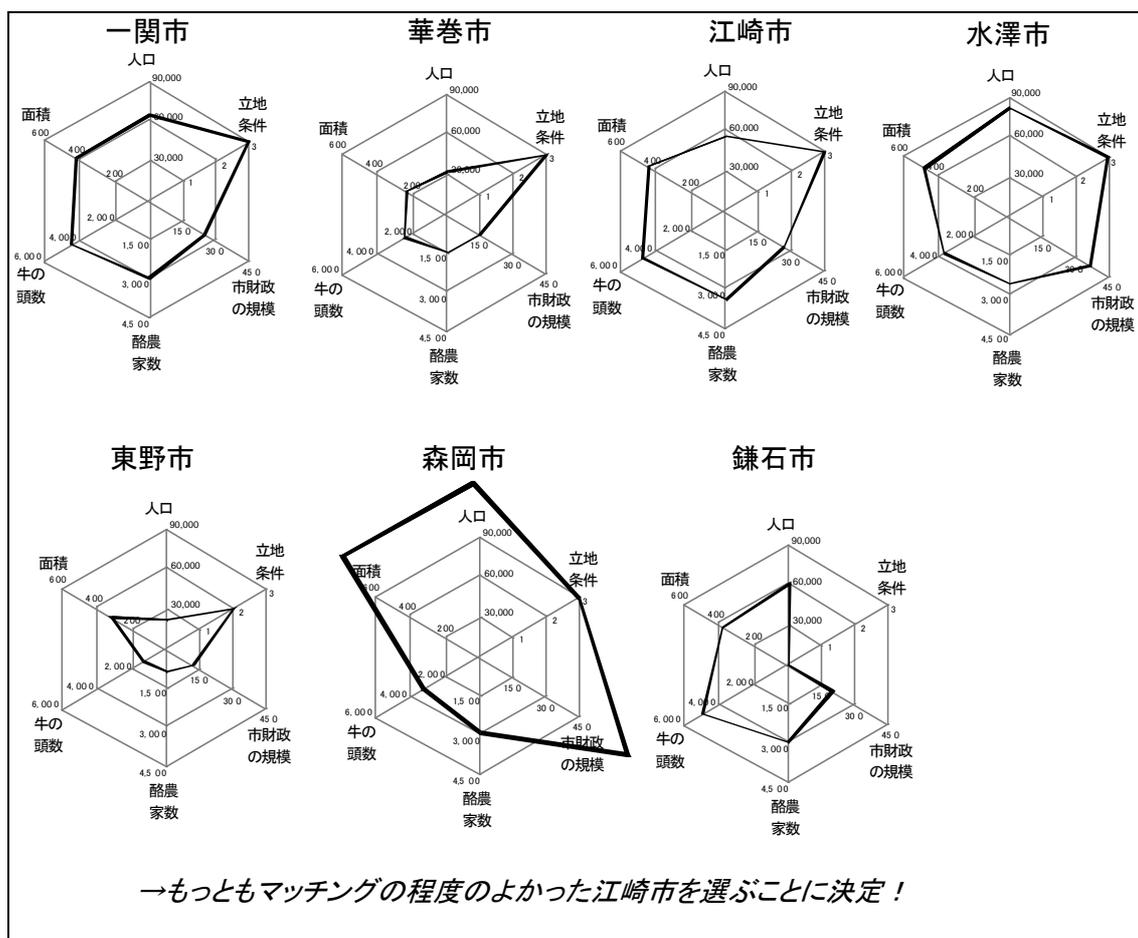
The screenshot shows a Microsoft Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J	K
1	1		58								
2	2		35								
3	3		5								
4	4		50								
5	5		42								
6	6		23								
7	7		33								
8	8		8								
9	9		48								
10	10		53								
11	11		11								
12	12		56								
13	13		19								
14	14		47								
15	15		39								
16	16		15								
17	17		37								
18	18		17								
19	19		22								
20	20		2								
21	21		33								
22	22		23								
23	23		46								
24	24		21								
25	25		1								
26	26		54								
27	27		46								
28	28		9								
29	29		40								
30	30		37								
31	31										
32	32										
33	33										

4.3 適用・比較グループの決定方法2：マッチングデザイン（地域マッチング）

実施地区がすでに決まっている場合に、複数の近隣地区からもっとも似た地区を選ぶ方法である。どれだけ似ている地区を選べるかがポイントとなる。人口規模、所得水準、ロケーション（川沿いか山沿いか）、市場や都市中心部からの距離などがマッチング指標としてよく用いられる。これらのマッチング指標を用いて各地区のレーダーチャートを描いて、一番似た地区を選ぶ方法が勧められる。なお、レーダーチャートの代わりに、各マッチング指標に関して点数をつけて合計点によって似た地区を選ぶ方法もよく用いられる。

その後、実施地区と比較地区からそれぞれ一定数のサンプル世帯を選んでインタビューを行うことになる⁴。



(出所) 佐々木亮(2002)『政策評価トレーニングブック』

⁴ なお、必ずしも実施グループと比較グループのサンプル数が同数である必要はなく、予算と時間の都合から、実施グループのサンプル数と比較グループのサンプル数は最大 3:1 までアンバランスであっても許容できると言われている。ただし、1 グループの最小限のサンプル数である 25 を満たすためには、90:30（あるいは 75:25）のサンプル数が必要ということになる。また、経験上、4:1 まで許容されるという意見もある。

ケーススタディ A: マイクロファイナンス事業 (RCT)

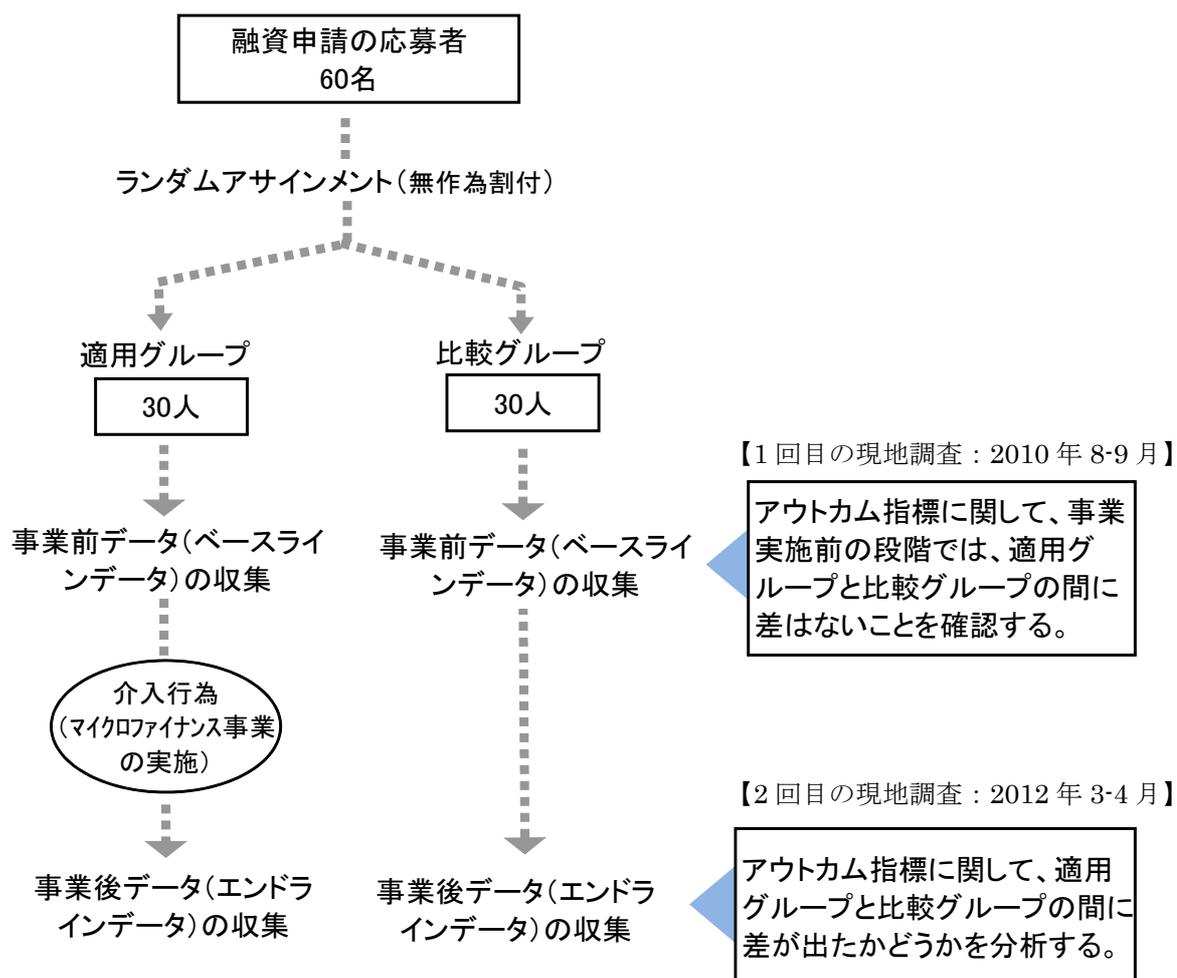
国際 NGO「国境なき隣人の会」は、次のとおりにアサインメントを実施した。

・ サンプルサイズの決定

資金（融資の原資）の制約から最大 30 名の農村女性に融資することにした。一方、融資を受けない同数の農村女性（30 名）にもインタビューを実施する。

・ 適用・比較グループの決定方法 1：ランダムアサインメント

マイクロファイナンスによる融資を希望する農村女性から融資申請を受け付ける。60 名まで応募者が集まったら募集を終了する。そして、コインの表と裏（実際にはエクセルの「サンプリング」(Sampling)を使用)で 30 名ずつのグループに分ける。表に当たった 30 名には研修に参加してもらい、研修を無事に終了した農村女性には融資を実行する。また、事業前および事業後のインタビューにも回答してもらおう。一方、裏に当たった 30 名には今年度の融資は適用されないが、来年度に優先的に適用することを伝えて、そのかわり今年度のインタビューに協力してもらおう。



ケーススタディ B: 地場野菜の共同出荷事業 (地域マッチング)

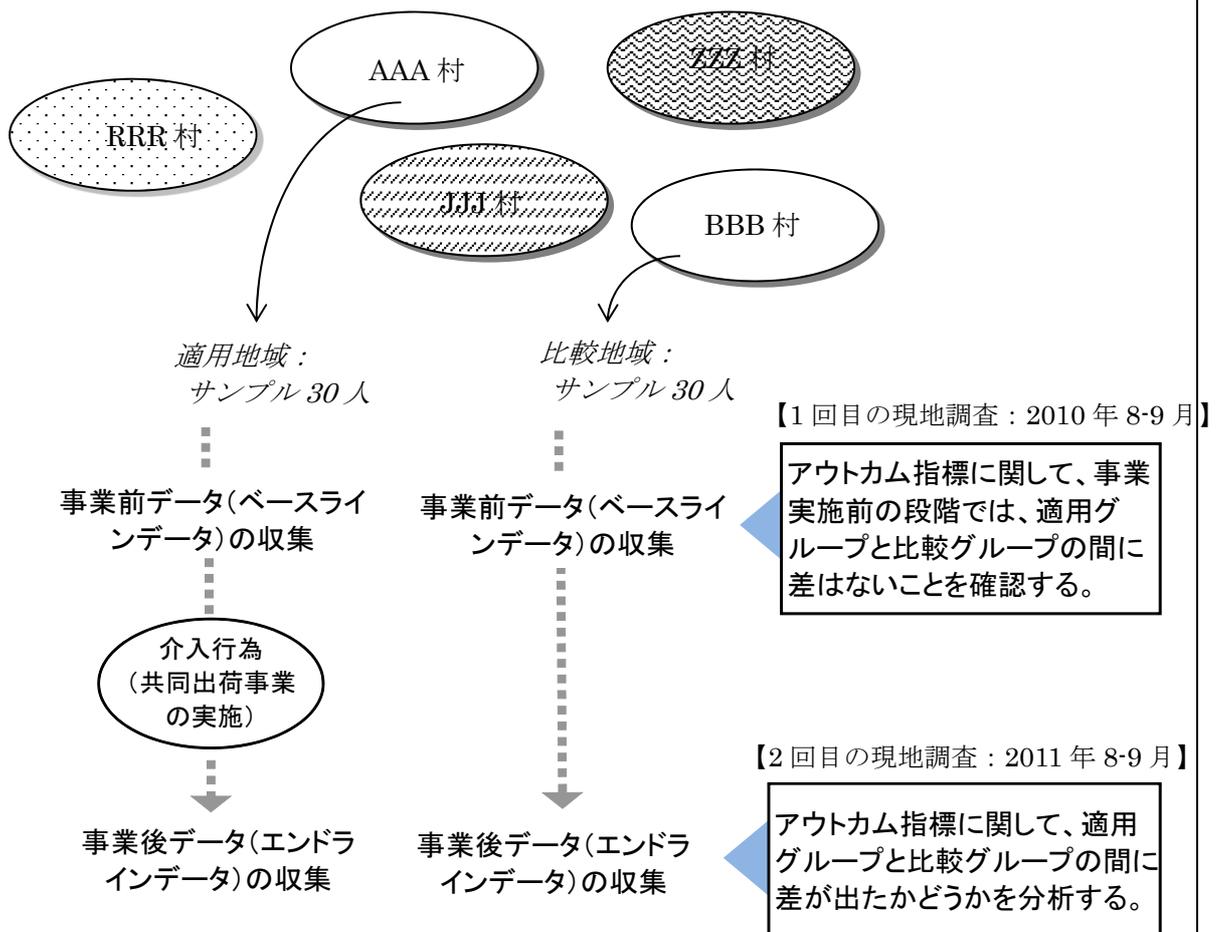
国際 NGO 「かるがも！」は、次のとおりの企画書をまとめた。

・サンプルサイズの決定

地場野菜の共同出荷事業に参加している世帯は約 100 世帯だが、そのうち 30 世帯をランダムサンプリングで選んでインタビューに応じてもらった。

・適用・比較グループの決定方法 2 : マッチングデザイン (地域マッチング)

地場野菜の共同出荷事業を実施している AAA 村と非常によく似た村をまずは選定する。具体的には周辺の村で、まだ共同出荷事業を実施していない村を 3~5 個リストアップして、その中から諸条件がもっとも近い BBB 村を比較対象として選定した (地域マッチング)。そして BBB 村からランダムサンプリングで 30 名を選定してインタビューに応じてもらった。



ステップ5：事業前データ（ベースラインデータ）の収集と分析

事業前データとは、事業実施前のアウトカム指標値のことである。事業実施前は、実施グループと比較グループのアウトカム指標値に違いがないことが理想である。

- ・ 質問票（クエスチオネア）の試験的適用（その後、微修正）
- ・ 質問票を使った訪問調査によるアウトプット指標値&アウトカム指標値の収集
- ・ 表計算ソフト（エクセル）への入力
- ・ 統計検定の実施：平均値などに差があるとは言えないことを確認する

5.1 質問票（クエスチオネア）の試験的適用（その後、微修正）

仮に作成した質問票を試験的に適用してみる。パートナーである現地 NGO の職員や現地の村役場の職員に答えてもらうとよい。また、実際に実施グループの世帯 3 件ほどに対して適用してみる。その後、得られた知見に基づいて微修正を行って完成させる。

5.2 質問票を使った訪問調査によるアウトプット指標値&アウトカム指標値の収集

完成した質問票に基づいて訪問調査を実施する。現場で採用した訪問調査者（サーベイヤ）に対する 1 日トレーニングを実施する。その後、十分な時間と旅費を与えて訪問調査を実施してもらう。

5.3 表計算ソフト（エクセル）への入力

エクセルにデータ入力を行う。エクセルシートの雛形は以下のとおり（別添 2 に拡大版を掲載している）。

ID 番号&グループ識別 アウトプット指標 アウトカム指標（事業前（2010）） アウトカム指標（事業後（2012））
 （事業後に収集して入力する）

ID	グループ (1=実施グループ、 0=比較グループ)	アウトプット指標		短期アウトカム(2010)			中期アウトカム(2010)		長期アウトカム(2010)		短期アウトカム(2012)		中期アウトカム(2012)		長期アウトカム(2012)	
		融資を受けたか? (1=はい、0=いいえ)	融資を受けた額 (ルピー)	事業の実施有無 (2010) (1=はい、 0=いいえ)	新規事業の業種 (2010) (カテゴリー1~4)	世帯の収入額 (2010) (ルピー/年)	世帯に意思決定者 (2010) (1=自分、2=配偶 者、3=その他)	世帯全体の 教育支出 (2010) (ルピー/年)	保健状況 (2010) (満足度0~4)	事業の実施有無 (2012) (1=はい、 0=いいえ)	新規事業の業種 (2012) (カテゴリー1~4)	世帯の収入額 (2012) (ルピー/年)	世帯に意思決定者 (2012) (1=自分、2=配偶 者、3=その他)	世帯全体の 教育支出 (2012) (ルピー/年)	保健状況 (2012) (満足度0~4)	
1	0	0	0	1	2	50,000	2									
2	1	1	100,000	0		75,000	1									
3	0	0	0	0		37,500	3									
4	0	0	0	0		12,500	3									
5	1	0	0	0		25,000	1									
6	0	0	0	0		12,500	2									
7	1	1	50,000	0		50,000	1									
8	1	1	80,000	0		62,500	1									
9	0	0	0	0		2,500	2									
10	1	0	0	0		20,000	2									
11	0	0	0	0		0	0									
12	0	0	0	0		0	0									
13	0	0	0	0		0	0									
14	0	0	0	0		0	0									

5.4 適用可能な場合には統計検定の実施：平均値などに差があるとは言えないことを確認

収集されたアウトカム指標（事業前）それぞれに関して、エクセルで「独立のt検定」を行う。別添3にその手順をまとめている。その結果を次のような表とグラフにまとめる。

表のまとめ方

表は以下のように表示する。平均値とサンプル数(n)は必須。さらに標準偏差を追加して表示することも勧められる。ただし、平均値やサンプルサイズと違って、標準偏差は一般読者には馴染みが薄く逆に混乱を招く可能性があることから、混乱の可能性がある場合には入れなくてもよい。照会があった場合には、平均値とサンプルサイズ以外の情報も提供することにする。

アウトカム： XXXX

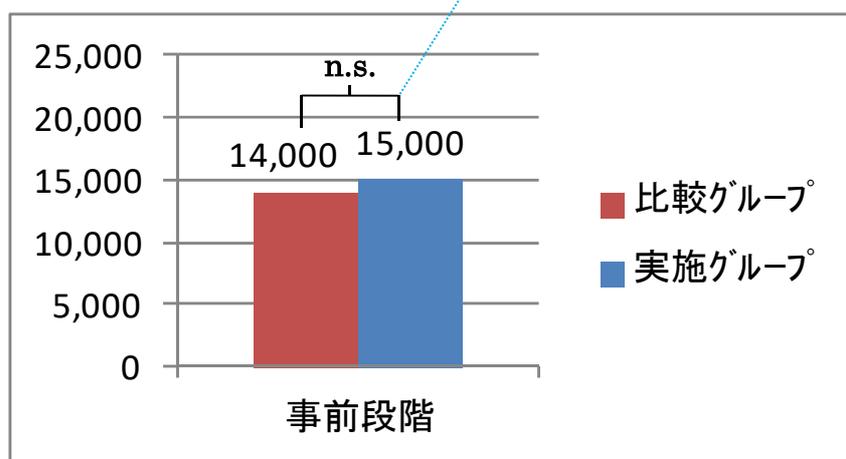
	事前段階
実施グループ	平均値 XXX (n=xx)
比較グループ	平均値 XXX (n=xx)
両グループの差	差 XXX ^{***} / ^{**} / [*] /n.s.

有意差検定 (t 検定) の結果を書き加える。
 ***1%水準で有意
 **5%水準で有意
 *10%水準で有意
 n.s.有意とは言えない。

グラフのまとめ方

グラフは以下のようにまとめる。エクセルで棒グラフを作り、有意差検定の結果を星印の個数で示す (***/**/*/n.s.)。

なお星印は、エクセルメニューから「テキストボックスを追加」を選んで手で書き加える。直線も手で図形描画から直線を選択して書き加える。



ケーススタディ A: マイクロファイナンス事業 (RCT)

国際 NGO「国境なき隣人の会」は、次のとおりに事業前データを収集した。

- ・ 質問票 (クエスチョネア) の試験的適用 (その後、微修正)
- ・ 質問票を使った訪問調査によるアウトカム指標値の収集

現地で雇用したネパール人の大学生 3 名を訪問調査者としてトレーニングして訪問調査を実施した。3 名×5 世帯×4 日間 (融資対象世帯 2 日間&非融資対象世帯 2 日間) =60 世帯から情報収集した。

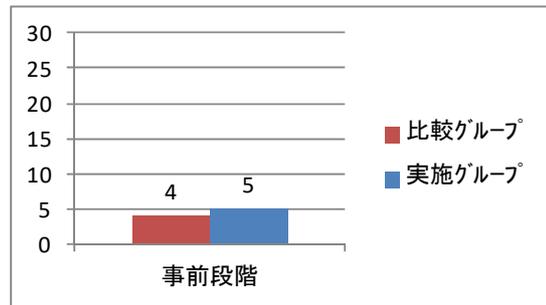
- ・ 統計検定の実施: アウトカム指標の平均値に差があるとは言えないことを確認した (偶然で起こり得る程度の差の範囲内であるということ)。

短期アウトカム: 新規事業が実施される。

指標: 事業の件数 (件) (既存事業の拡張を含む)

	事前段階
実施グループ ⁵	5 (n=30)
比較グループ	4 (n=30)
平均値の差	差 +1

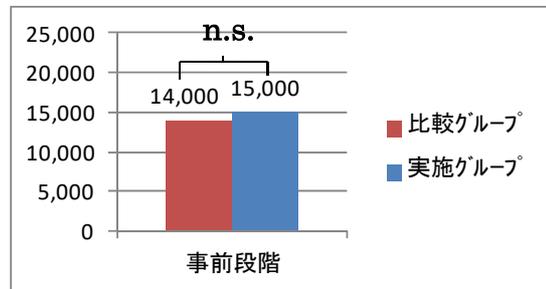
(注) 件数なので t 検定は適用できない⁵。



中期アウトカム: 農村女性の収入が増加する。

指標: 世帯の収入額 (ルピー/年間)

	事前段階
実施グループ ⁵	15,000 (n=30)
比較グループ	14,000 (n=30)
平均値の差	+1,000 N.S.

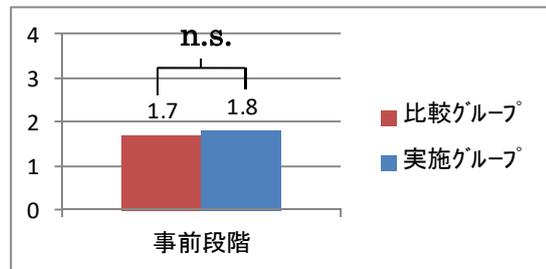


長期期アウトカム: 農村女性の地位向上。

指標: 女性が主たる意思決定者であるという認識 (5 段階 5~1)

	事前段階
実施グループ ⁵	1.8 (n=30)
比較グループ	1.7 (n=30)
平均値の差	0.1 N.S.

* 世帯全体の教育・保健状況に関する表・グラフは省略した。



⁵ 比率を検定するカイ二乗検定なら適用できるが本書では省略。

ケーススタディ B: 地場野菜の共同出荷事業（地域マッチング）

国際 NGO「かるがも！」は、次のとおりに事業前データを収集した。

- ・ 質問票（クエスチョネア）の試験的適用（その後、微修正）
- ・ 質問票を使った訪問調査によるアウトカム指標値の収集

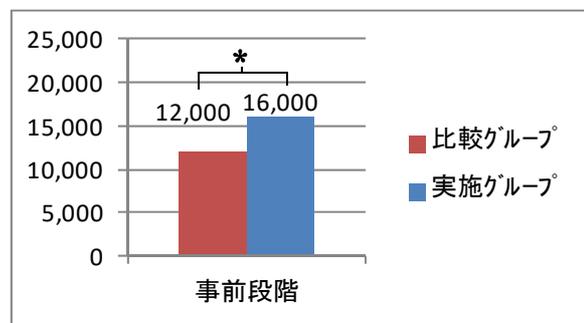
現地で雇用したカンボジア人の大学生 3 名を訪問調査者としてトレーニングして訪問調査を実施した。3 名×5 世帯×4 日間（事業対象村 2 日間&非事業対象村 2 日間）＝60 世帯から情報収集した。

- ・ 統計検定の実施：アウトカム指標の平均値に差があるとは言えないことを確認しようとしたが、結果は、有意な差が確認されてしまった。マッチングが適切に機能しなかったと言わざるを得ない。仕方ないので、事業後データの分析の際に、「差の差の計算」（二重引き算法）を用いることで補正することとした。

短期アウトカム 1：地場野菜の売上増加

指標：売上高／世帯（半年間）（円換算）

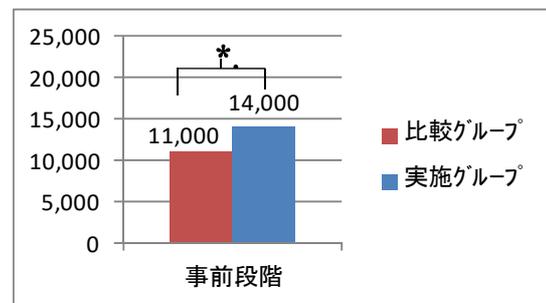
	事前段階
実施グループ	16,000 (n=30)
比較グループ	12,000 (n=30)
平均値の差	+4,000*



短期アウトカム 2：農民の収入増加

指標：収入額／世帯（半年間）（円換算）

	事前段階
実施グループ	14,000 (n=30)
比較グループ	11,000 (n=30)
平均値の差	+3,000*

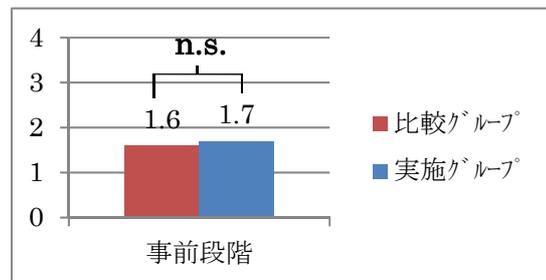


* 耐久消費財の有無の表とグラフは省略した。

長期期アウトカム：生活の改善

指標：生活の満足度（5段階）

	事前段階
実施グループ	1.7 (n=30)
比較グループ	1.6 (n=30)
平均値の差	0.1 n.s.



* 世帯全体の教育・保健状況に関する表・グラフは省略した。

ステップ6：事業実施およびそのモニタリング

事業前データ（ベースラインデータ）の収集が終わったら事業を実施する。また実施過程のモニタリングを行う。特に以下の点をチェックする。

- ・投入は計画どおりだったか（人員、資金、物品）
- ・活動の回数・規模は計画どおりだったか
- ・アウトプットが計画通りに実現したか

6.1 投入は計画どおりだったか（人員、資金、物品）

投入（人員、資金、物品）の投入は計画どおりだったかどうかをモニタリングする。

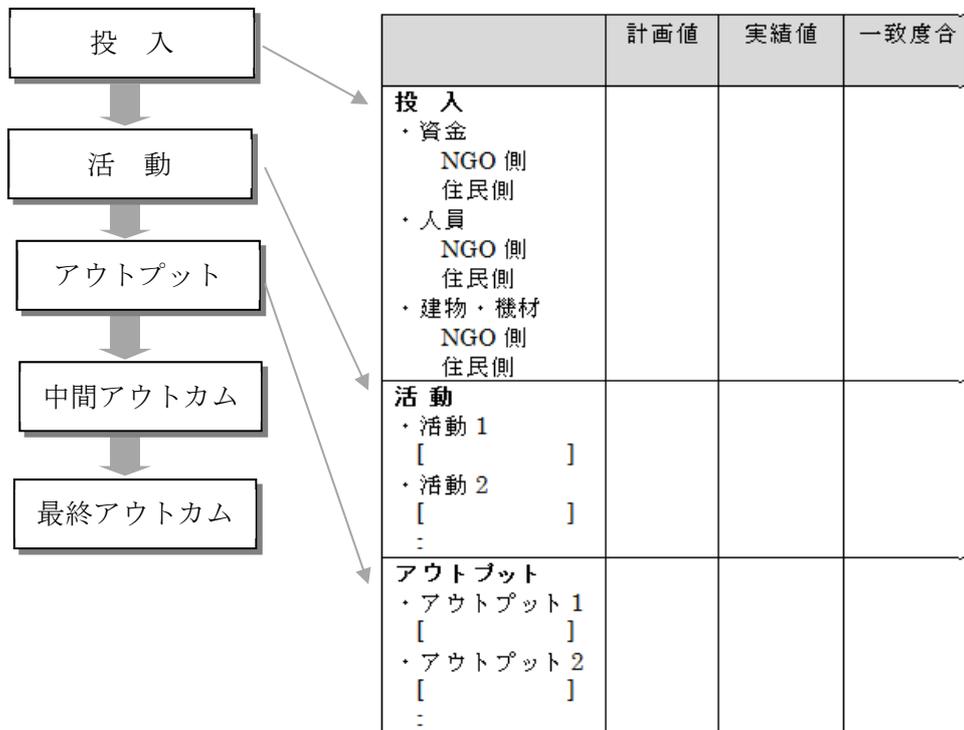
6.2 活動の回数・規模は計画どおりだったか

活動の回数・規模の投入は計画どおりだったかどうかをモニタリングする。

6.3 アウトプットが計画通りに実現したか

アウトプットの算出量（人員の場合は人数）の投入は計画どおりだったかどうかをモニタリングする。

いずれのレベルもロジックモデルに従い、以下のような対比一覧表にまとめるとよい。



ケーススタディ A: マイクロファイナンス事業 (RCT)

国際 NGO「国境なき隣人の会」によるネパールの対象地域での投入（人員、資金、事務所スペース）、活動の回数・規模は当初計画どおりであった。無事に 60 名の農村女性から申請があり、ランダムアサインメントによってそれぞれ 30 名のグループに分けた。そして当初計画どおりの活動が実施された。その結果、アウトプットである「融資を受けた女性」に関して計画どおりのアウトプットが生み出されたと言える。

	計画値	実績値	一致度合
投入			
・資金	XX	XX	XX
・人員	XX	XX	XX
・建物・機材	XX	XX	XX
活動			
・活動 1[農村女性向けの研修]	XX	XX	XX
・活動 2[融資申請受付・審査事務]	XX	XX	XX
・活動 3[貸出実行・返済受付事務]	XX	XX	XX
アウトプット			
・アウトプット 1 [融資を受けた女性]	30 人	28 人	ほぼ一致

ケーススタディ B: 地場野菜の共同出荷事業 (地域マッチング)

国際 NGO「かるがも！」によるカンボジアの対象農村での投入（人員、資金、事務所スペース）、活動の回数・規模は当初計画どおりであった。

	計画値	実績値	一致度合
投入			
・資金	XX	XX	XX
・人員	XX	XX	XX
・建物・機材	XX	XX	XX
活動			
・活動 1[農民性向けの研修]	XX	XX	XX
・活動 2[小型トラック借受]	XX	XX	XX
・活動 3[共同出荷の実行]	XX	XX	XX
アウトプット			
・アウトプット 1 [共同出荷の実施実績]	月 2 回 半年 12 回	月 2 回 半年 12 回	一致

ステップ7: 事業後データ (エンドラインデータ) の収集と分析

事業後データとは、事業実施後のアウトカム指標値のことである。事業実施後は、実施グループと比較グループのアウトカム指標値に違いがあることが理想である。

- ・ 質問票 (クエスチョネア) を使った訪問調査によるアウトカム指標値の収集
- ・ 統計検定の実施: 平均値に差があると言えることを確認する

7.1 質問票 (クエスチョネア) を使った訪問調査によるアウトカム指標値の収集

事業前データの収集の際に雇用した訪問調査者 (サーベイヤー) を再度雇用して、十分な時間と旅費を与えて訪問調査を実施してもらう。

7.2 統計検定の実施: 平均値に差があると言えることを確認する

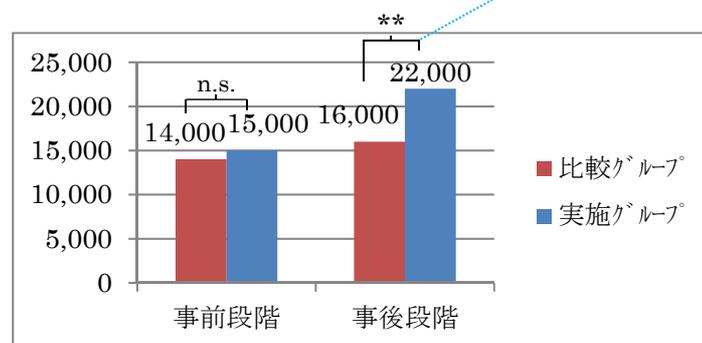
表のまとめ方

実施グループと比較グループでt検定を行い、次のような表とグラフにまとめる。

	事前段階	事後段階
実施グループ	平均値 XXX (n=xx)	平均値 XXX (n=xx)
比較グループ	平均値 XXX (n=xx)	平均値 XXX (n=xx)
両グループの差	差 XXX ($p=xx$)	差 XXX ($p=xx$)

グラフのまとめ方

グラフは以下のようにまとめる。エクセルで棒グラフを作り、有意差検定の結果を星印の個数で示す (***/**/* / n.s.)。なお星印は、エクセルメニューから「テキストボックスを追加」を選んで手で書き加える。直線も手で図形描画から直線を選択して書き加える。



ケーススタディ A: マイクロファイナンス事業 (RCT)

国際 NGO「国境なき隣人の会」は、次のとおりに事業後データを収集した。

- ・ 質問票を使った訪問調査によるアウトカム指標値の収集

現地で雇用したネパール人の大学生 3 名を再度雇用して、事業前データのとおりと同じ 30 世帯×2 グループ=60 世帯へインタビューを実施した。

- ・ 統計検定の実施

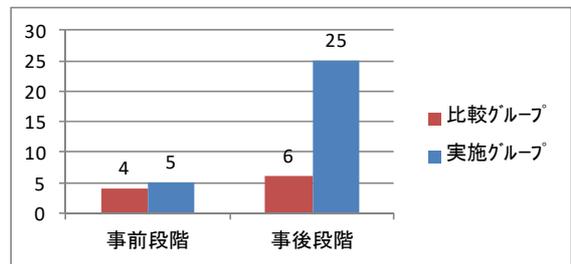
事業後データを使用して、短期アウトカム指標の平均値に差があったことを確認した(+19 件)。また、中期アウトカム指標の平均値にも差があったことを確認した(+6,000 ルピー)。さらに、長期アウトカム指標の平均値にもすでに差が確認された(+0.7)。

短期アウトカム：新規事業が実施される

指標：事業の件数 (件) (既存事業の拡張を含む)

	事前段階	事後段階
実施グループ ⁶	5 (n=30)	25 (n=30)
比較グループ ⁶	4 (n=30)	6 (n=30)
平均値の差	差 +1	差 +19

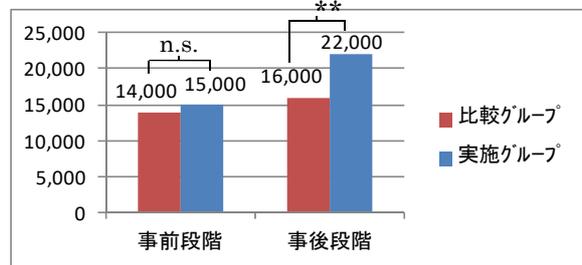
(注) 件数なので t 検定は適用できない⁶。



中期アウトカム：農村女性の収入が増加する

指標：世帯の収入額 (ルピー/年間)

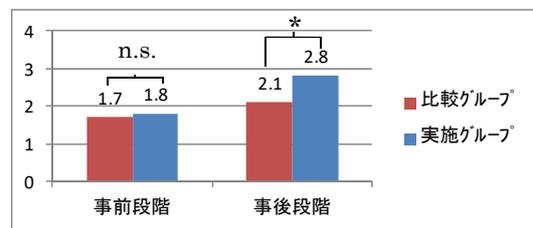
	事前段階	事後段階
実施グループ ⁶	15,000 (n=30)	22,000 (n=30)
比較グループ ⁶	14,000 (n=30)	16,000 (n=30)
平均値の差	+1,000 N.S.	+6,000**



長期アウトカム：農村女性の地位向上

指標：女性が主たる意思決定者であるという認識 (5 段階 s~e)

	事前段階	事後段階
実施グループ ⁶	1.8 (n=30)	2.8 (n=30)
比較グループ ⁶	1.7 (n=30)	2.1 (n=30)
平均値の差	+0.1 N.S.	+0.7*



* 世帯全体の教育・保健状況に関する表・グラフは省略した。

⁶ 比率を検定するカイ二乗検定なら適用できるが本書では省略。なお、この短期アウトカムでは目で見ても十分に大きな差が現れているので、統計検定を実施しなくてもさしたる問題にはならない。

ケーススタディ B: 地場野菜の共同出荷事業 (地域マッチング)

国際 NGO「かるがも！」は、次のとおりに事業後データを収集した。

- ・質問票を使った訪問調査によるアウトカム指標値の収集

現地で雇用したカンボジア人の大学生 3 名を再度雇用して、事業前データのとくと同じ 30 世帯×2 グループ=60 世帯へインタビューを実施した。すでに引っ越していた世帯もあり最終的に 25 世帯×2 グループ=50 世帯から回答を得られた。

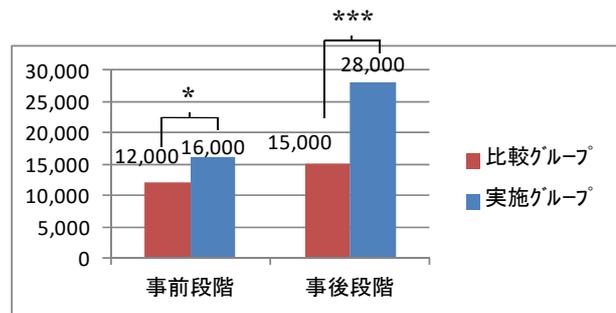
- ・統計検定の実施

2 重引き算法により、短期アウトカム 1 の平均値に差があったことを確認した (+9,000 円 (=13,000 円 - 4,000 円))。また、短期アウトカム 2 の指標の平均値にも差があったことを確認した (+6,000 円 (=9,000 円 - 4,000 円))。このように事前段階でも平均値差が有意であった (偶然で起こり得る程度よりも大きな差があったということ) 場合には、事後の平均値差から事前の平均値差を引き算して、インパクトを算出する。長期アウトカムの指標の平均値にもすでに差が確認された (+0.6)。

短期アウトカム 1: 地場野菜の売上増加

指標: 売上高 / 世帯 (半年間) (円換算)

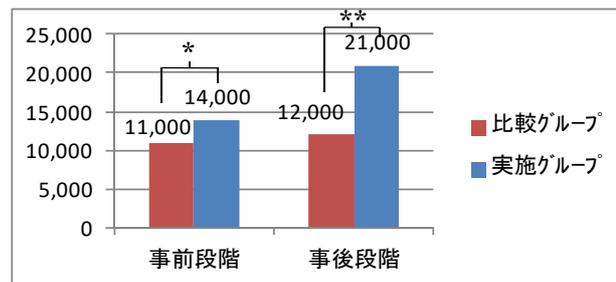
	事前段階	事後段階
実施グループ	16,000 (n=30)	28,000 (n=25)
比較グループ	12,000 (n=30)	15,000 (n=25)
平均値の差	+4,000*	+13,000***



短期アウトカム 2: 農民の収入増加

指標: 収入額 / 世帯 (半年間) (円換算)

	事前段階	事後段階
実施グループ	14,000 (n=30)	21,000 (n=25)
比較グループ	11,000 (n=30)	12,000 (n=25)
平均値の差	+3,000*	+9,000**

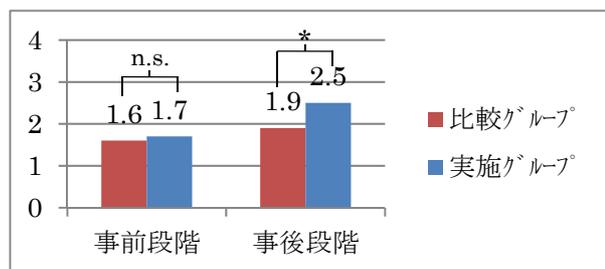


*耐久消費財の有無の表とグラフは省略した。

長期期アウトカム 2: 生活の改善

指標: 生活の満足度 (5 段階)

	事前段階	事後段階
実施グループ	1.7 (n=30)	2.5 (n=25)
比較グループ	1.6 (n=30)	1.9 (n=25)
平均値の差	+0.1 n.s.	+0.6*



*世帯全体の教育・保健状況に関する表・グラフは省略した。

ステップ8：インパクト評価結果の確定

ステップ5～7で得たデータを使用してインパクト評価結果を確定する。

- ・ 収集されたデータを分かりやすく表とグラフにして示す。
- ・ 評価枠組に基づいて評価結果を決定する。

8.1 収集されたデータを分かりやすく表とグラフにして示す

すでにステップ7で実施しているとおりだが、表とグラフを資金支援者および会員が容易に理解できるように整備する。

8.2 評価枠組に基づいて評価結果を決定する

インパクト評価の対象とした事業が効果（インパクト）があったのかなかったのかを文章で明記する。またステップ7で作成した表やグラフを装入することが勧められる。

ケーススタディ A: マイクロファイナンス事業（RCT）

国際 NGO「国境なき隣人の会」は、次のとおりにインパクト評価結果とそれに基づく提言をとりまとめた。

『マイクロファイナンス事業の実施によって、融資を受けた農村女性の多くが新規事業を開始したことが分かった（+19 件⁷）。また、収入が増加したことが分かった（+6,000 ルピー）。さらに、早くも長期的効果として考えていた世帯における女性の地位の向上と教育・保健状況の改善もある程度実現していることが分かった（+0.7 段階）。厳格なインパクト評価によって、マイクロファイナンス事業の有効性が確かめられた。その結果は満足できるものだった。したがって、来年度の事業として対象地域を拡大すべきである。』

ケーススタディ B: 地場野菜の共同出荷事業（地域マッチング）

国際 NGO「かるがも！」は、次のとおり次のとおりにインパクト評価結果をとりまとめた。

『地場野菜の共同出荷事業によって、対象地域の農民の農産品売上が増加して（+9,000 円）、世帯収入が増加していることが分かった（+6,000 円）。また、早くも長期的効果として考えていた生活の満足度がある程度向上していることも分かった（+0.6 段階）。今回の厳格なインパクト評価の実施によって、「かるがも！」が長年支援・実施してきた同事業の有効性を確かめることができた。その結果は満足できるものだった。したがって、支援者各位にさらなる支援をお願いしたい。』

⁷ 事前段階ですでに 1 件の差があったので、 $19-1=18$ 件とするとより正確である。

ステップ9：インパクト評価報告書の作成

今までの作業の結果をとりまとめて評価報告書を作成する。評価報告書の目次案は次のとおりである。「要約」は、評価結果がすぐに分かるように1ページ、あるいは最長でも2ページでとりまとめる。

評価報告書 目次案

謝辞

要約

1. 対象 NGO の概要
 2. 対象事業の背景・目的・概要
 3. 対象地域・対象人口
 4. 評価枠組（評価項目と評価基準で構成）
 5. ロジックモデル（アウトプット指標、アウトカム指標も明示）
 6. データ収集方法
 7. インパクト評価結果
 - 4-1. 事業前指標値（ベースラインデータ）
 - 4-2. 事業実施とモニタリング結果
 - 4-3. 事業後指標値（エンドラインデータ）
 - 4-4. インパクト評価結果
 8. 提言と教訓
 9. 本件評価の制約
- 別添：収集資料リスト、インタビュー結果の概要、サーベイ結果の概要

報告書が完成したら報告会を開催する。

ケーススタディ A： マイクロファイナンス事業（RCT）

国際 NGO「国境なき隣人の会」はインパクト評価報告書（ドラフト）をとりまとめて、日本の本部事務所で報告会を開催した。報告会でのコメントを踏まえて報告書は最終化され、ハードコピーが関係者に配布されたほか、同 NGO のウェブサイトでも公開された。

ケーススタディ B： 地場野菜の共同出荷事業（地域マッチング）

国際 NGO「かるがも！」は、はインパクト評価報告書（ドラフト）をとりまとめて、日本の本部事務所で報告会を開催した。報告会でのコメントを踏まえて報告書は最終化され、ハードコピーが関係者に配布されたほか、同 NGO のウェブサイトでも公開された。

ステップ 10：意思決定への利用

インパクト評価の結果を意思決定に利用する。

- ・パイロット事業として実施した場合：本格的実施をするかどうかの意思決定
- ・説明責任遂行の一環として実施した場合：資金支援者への説明と事業を継続するかどうかの意思決定

10.1 パイロット事業として実施した場合：本格的実施をするかどうかの意思決定

パイロット事業の場合には、インパクト評価の結果を踏まえて、本格実施に移行するかどうかを決定する。なお、インパクト評価の結果は、重要ではあっても意思決定のための情報の一部に留まるわけで、その他の情報、例えば次年度以降の資金支援獲得額、人材投入の優先順位、組織としての中長期計画なども総合的に勘案して意思決定がなされるべきである。よく言われるように、「評価は意思決定に取って代わるわけではない」ということである。

ケーススタディ A: マイクロファイナンス事業 (RCT)

国際 NGO「国境なき隣人の会」(仮称)は理事会を開催した。そして、インパクト評価報告書の提言を受け入れて、マイクロファイナンスを新しい活動の柱として適用していくことにした。まずはネパールのシズリ県で適用する村を周辺に拡大していくことを次年度計画に盛り込むことになった。

10.2 説明責任遂行の一環として実施した場合：資金支援者への説明と事業を継続するかどうかの意思決定

説明責任のために実施した場合には、インパクト評価の結果を広く広報することに注力すべきである。評価が終わるとそれで組織として満足してしまうことがあるが、その後の広報活動が重要である。お金をかけて綺麗な報告書として多めに印刷製本して、現在及び将来の資金支援者に積極的に配布する。また、ウェブ上で公開することも検討する。さらに、広報セミナーの開催も重要である。加えて、事業継続の意思決定にも必要に応じて評価結果を利用する。

ケーススタディ B: 地場野菜の共同出荷事業 (地域マッチング)

国際 NGO「かるがも！」(仮称)は、はインパクト評価報告書(ドラフト)をもとに、新たな広報活動を開始することになった。また、このインパクト評価結果を利用して、公益法人や行政機関からの補助金を得るための営業に力を注ぐことになった。

別添1：質問票（雛形）

以下はマイクロファイナンス事業に関する質問票の雛形である。基本は、ロジックモデルに沿って、活動、アウトプット、アウトカム(短期、中期、長期)に沿った質問を設定することである。

訪問調査の趣旨説明

1. 事業の活動への参加に関して

- 農村女性向けの研修への参加の有無（1:はい, 0:いいえ, 2 無回答/その他の回答）
- 融資説明会への参加の有無（1:はい, 0:いいえ, 2 無回答/その他の回答）

2. アウトプットに関して

- 融資を受けたかどうか（1:はい, 0:いいえ, 2 無回答/その他の回答）
（「1:はい」の場合、以下も質問する）
- 融資を受けたのはいつか？（MM/YYYY）
- 融資金額はいくらだったか？（XXX ルピー）

3. 短期アウトカム

- 新規事業を開始したか？（1:はい, 0:いいえ, 2 無回答/その他の回答）
（「1:はい」の場合、以下も質問する）
- 業種はなにか？（カテゴリー1～5 から選択する）
- 資本金はいくらか？（XXX ルピー）

4. 中期アウトカム

- 世帯全体の収入（年間）はいくらか？（XXX ルピー）
（あるいは項目別の月間支出額を質問して合計してもほぼ同様の情報が得られる）

5. 長期アウトカム

- 世帯において、主たる意思決定者は女性か？（5段階レーティング）
[4:おおいにそう思う, 3:そう思う, 2 どちらとも言えない, 1:そうは思わない, 0:全くそう思わない]
- 世帯全体の教育支出（月間）はいくらか？（XXX ルピー）
- 世帯全体の年間の保健支出（月間）はいくらか？（XXX ルピー）

6. 事業への全体的コメント及び改善に向けたアドバイス（オープンクエスチョン）

7. 属性に関する情報

- 名前
- 性別、年齢、民族（回答して差し支えない場合）
- 家族構成
- 連絡先（住所と携帯電話の番号）

御礼及び記念品の進呈

別添2：エクセル入力シート（雛形）

収集したデータは次のようにエクセルシートに入力して保存する。事業前（ベースライン）データ、事業後（エンドライン）データが1枚のシートに入力してあることが必須である。

ID 番号&グループ識別 アウトプット指標 アウトカム指標（事業前（2010）） アウトカム指標（事業後（2012））

	A	B	C		D		E		F		G		H		I		J		K		L		M		N		O		P	
1	ID	グループ (1=実施グループ、 0=比較グループ)	アウトプット指標		短期アウトカム(2010)						中期アウトカム(2010)						長期アウトカム(2010)						短期アウトカム(2012)		中期アウトカム(2012)		長期アウトカム(2012)			
			融資を受けたか? (1=はい、0=いいえ)	融資を受けた額 (ルピー)	事業の実施有無 (2010) (1=はい、 0=いいえ)	新規事業の業種 (2010) (カテゴリー1~4)	世帯の収入額 (2010) (ルピー/年)	世帯に意思決定者 (2010) (1=自分、2=配偶 者、3=その他)	世帯全体の 教育支出 (2010) (ルピー/年)	保健状況 (2010) (満足度0~3)	事業の実施有無 (2012) (1=はい、 0=いいえ)	新規事業の業種 (2012) (カテゴリー1~4)	世帯の収入額 (2012) (ルピー/年)	世帯に意思決定者 (2012) (1=自分、2=配偶 者、3=その他)	世帯全体の 教育支出 (2012) (ルピー/年)	保健状況 (2012) (満足度0~4)														
2	1	0	0	0	1	2	50,000	2	6,250	2	1	2	200,000	2	40,000	3														
3	2	1	1	100,000	0		75,000	1	9,375	3	1	3	300,000	1	60,000	4														
4	3	0	0	0	0		37,500	3	4,688	0	0	150,000	3	30,000	2															
5	4	0	0	0	0		12,500	3	1,563	1	0	50,000	3	10,000	2															
6	5	1	0	0	0		25,000	1	3,125	0	1	100,000	1	20,000	1															
7	6	0	0	0	0		12,500	2	1,563	1	0	50,000	2	10,000	2															
8	7	1	1	50,000	0		50,000	1	6,250	2	0	200,000	1	40,000	3															
9	8	1	1	80,000	0		62,500	1	7,813	3	1	250,000	1	50,000	4															
10	9	0	0	0	0		2,500	2	0	2	0	10,000	2	2,000	3															
11	10	1	0	0	0		20,000	2	2,500	1	0	80,000	2	16,000	2															
12	:	:	:	:	:		:	:	:	:	:	:	:	:	:															
13	:	:	:	:	:		:	:	:	:	:	:	:	:	:															
14	:	:	:	:	:		:	:	:	:	:	:	:	:	:															

別添3：エクセル操作法（等分散の検定、2群の有意差検定）

等分散の検定

2群の有意差検定をする前に等分散の検定を行う必要がある（「等分散のF検定」という）。以下のとおり操作する。

データの入力

以下のとおりにデータを入力する。

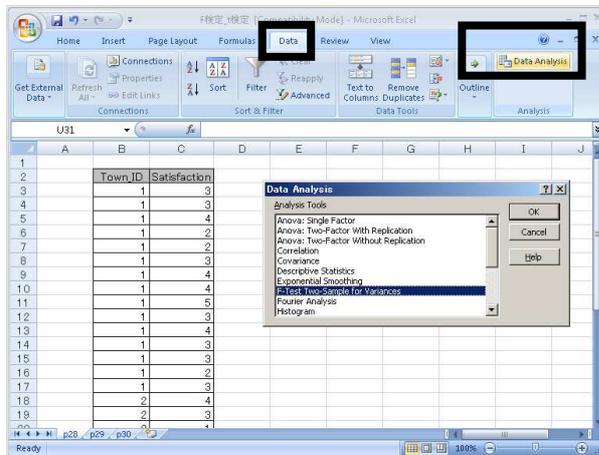
“適用グループ”（我校）（School_ID 1）

“比較グループ” 近隣校（School_ID 2）

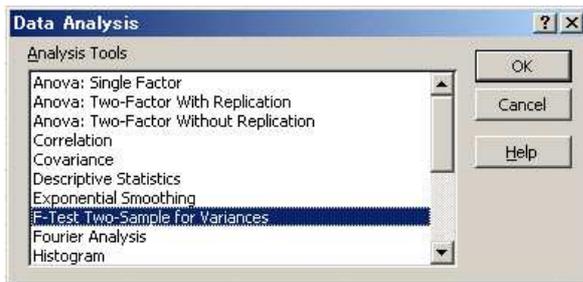
	A	B	C	D	E
1					
2		School_ID	Satisfaction		
3		1	3		
4		1	3		
5		1	4		
6		1	2		
7		1	2		
8		1	3		
9		1	4		
10		1	4		
11		1	5		
12		1	3		
13		1	4		
14		1	3		
15		1	3		
16		1	2		
17		1	3		
18		2	4		
19		2	3		
20		2	1		
21		2	3		
22		2	2		
23		2	3		
24		2	2		
25		2	3		
26		2	1		
27		2	3		
28		2	1		
29		2	2		
30		2	3		
31		2	2		
32		2	3		
33					

「データ分析」の選択

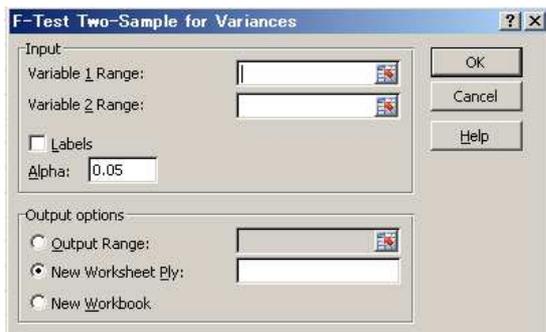
- データ > データ分析 を選択する。



- 「2群の分散のF検定」を選択する(F-Test: Two-sample for Variance)。

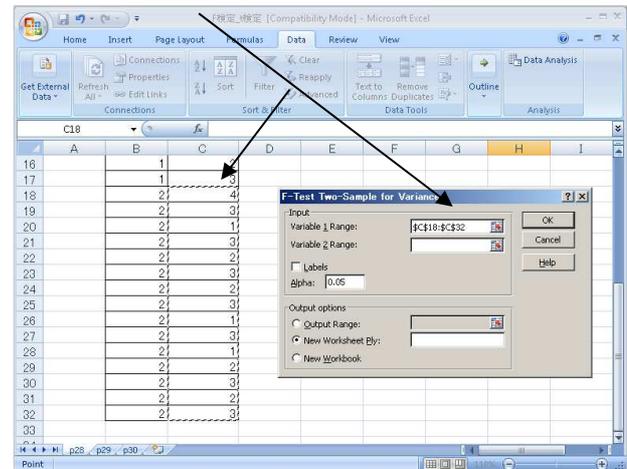


- 以下のボックスが現れる。

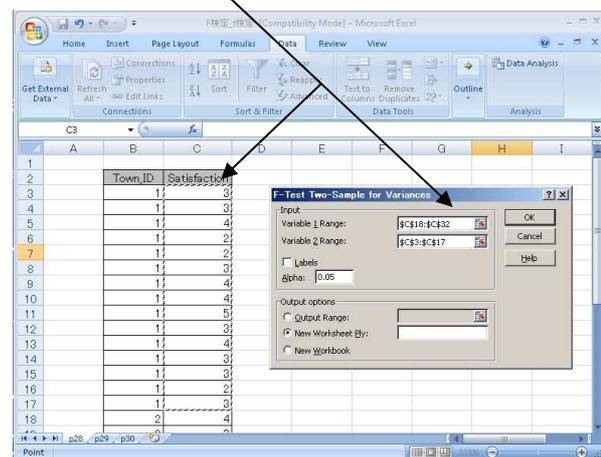


我校 (School_ID1) と 近隣校 (School_ID2) のどちらからデータを選択しても結果 (p値) は同じになるが、理解のしやすさのために、分散(=標準偏差の2乗)の大きい方を変数1に、分散(=標準偏差の2乗)の小さい方を変数2に選択する。

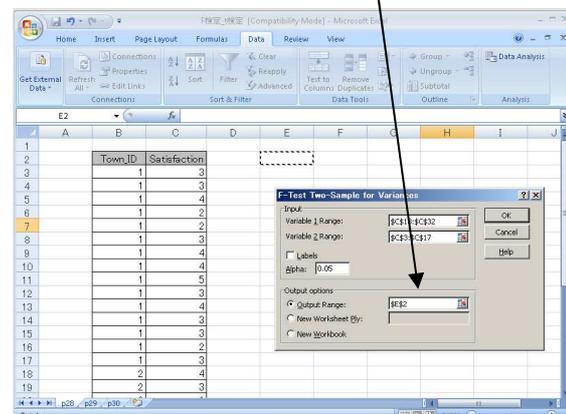
- 近隣校 (School_ID 2) のデータを選ぶ。



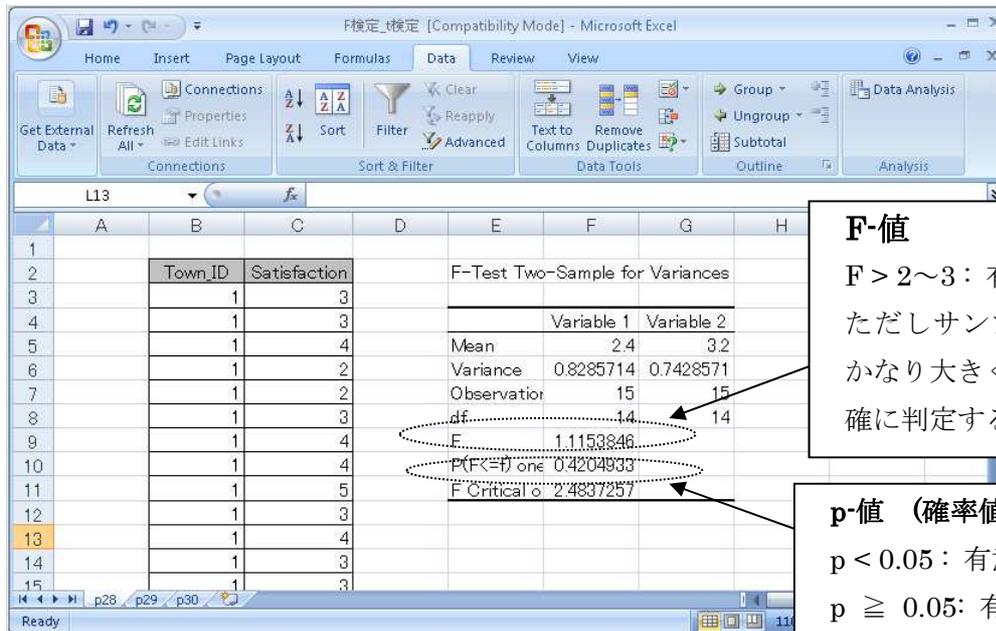
- 我校 (School_ID 1) のデータを選ぶ。



- 出力オプション > 出力範囲を選択して、空いている適当なセルを選ぶ。



分析結果を得る。



結論

“適用グループ” (我校) の分散と“比較グループ” (近隣校) の分散の比率は統計学的に有意であるとは言えない ($p=0.42 > 0.05$ (5%)). したがって、二つのデータは等分散であると仮定できる。

一方、p が 0.05 未満となった場合には次のとおりの結論になる。

結論

“適用グループ” (我校) の分散と“比較グループ” (近隣校) の分散の比率は統計学的に有意である ($p < 0.05$ (5%)). したがって、二つのデータは等分散であるとは言えない (「異分散」とも言う)。

【発展学習 : F 値の手計算】

F-値は、単純に、二つの分散の比率である。

$$F = \frac{\text{標準偏差}_{大}^2}{\text{標準偏差}_{小}^2} = \frac{\text{分散}_{大}}{\text{分散}_{小}} \quad \text{上記の例だと} \dots$$

	Variable 1	Variable 2
Mean	2.4	3.2
Variance	0.8285714	0.7428571
Observation	15	15

$$F = \frac{\text{分散}_{大}}{\text{分散}_{小}} = \frac{0.8285714}{0.7428571} = 1.1153846$$

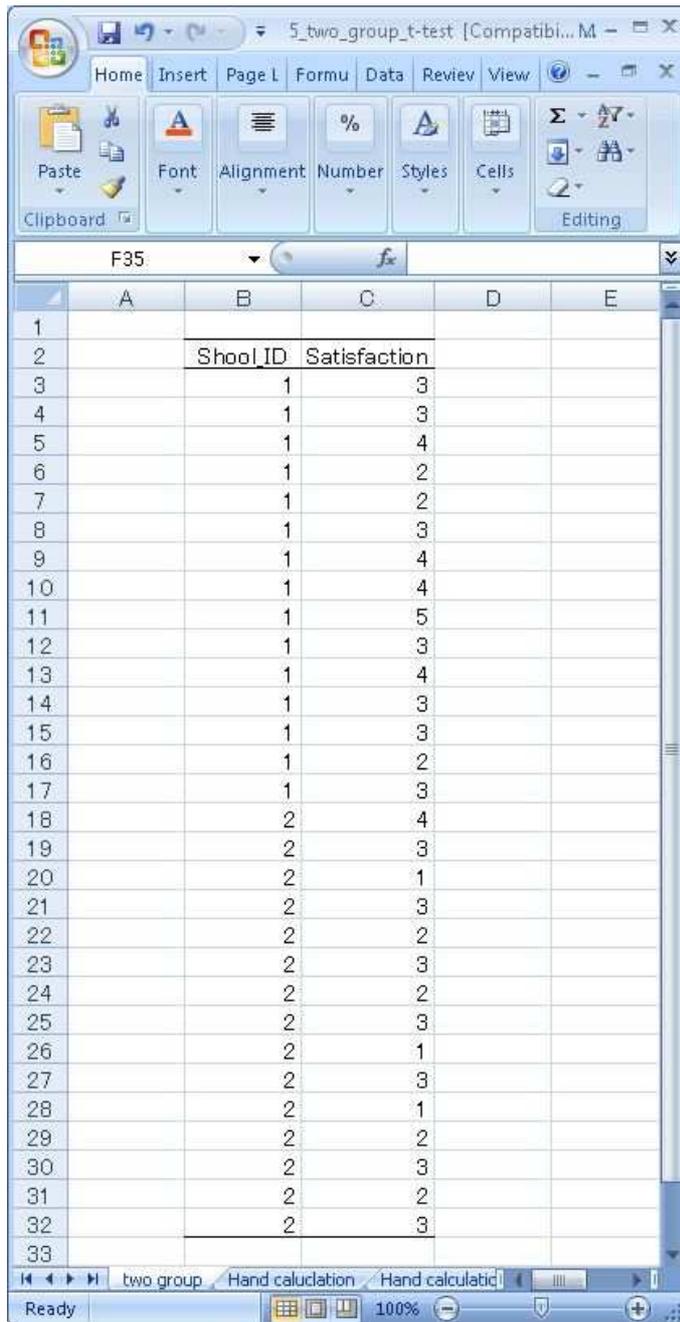
2群の有意差検定

等分散が確認できたら、2群の有意差検定を行う（「2群のt検定」、「独立のt検定」という）。以下のとおり操作する。

データの入力

“適用グループ”（我校）（School_ID 1）

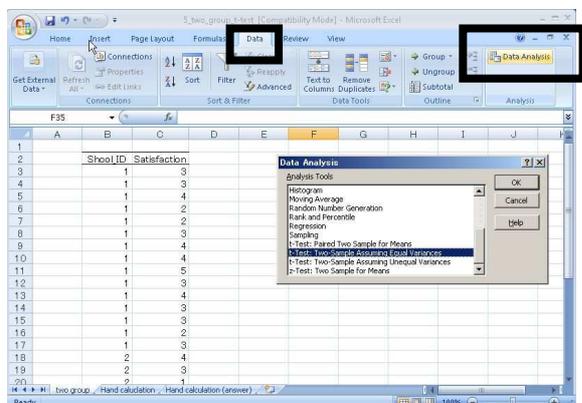
“比較グループ” 近隣校（School_ID 2）



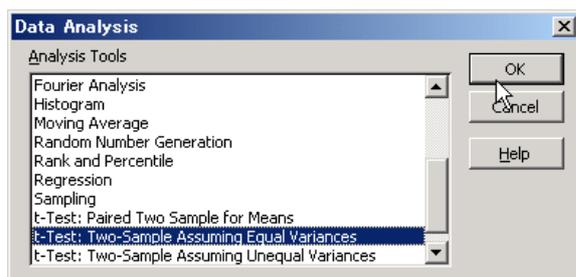
	A	B	C	D	E
1					
2		School_ID	Satisfaction		
3		1	3		
4		1	3		
5		1	4		
6		1	2		
7		1	2		
8		1	3		
9		1	4		
10		1	4		
11		1	5		
12		1	3		
13		1	4		
14		1	3		
15		1	3		
16		1	2		
17		1	3		
18		2	4		
19		2	3		
20		2	1		
21		2	3		
22		2	2		
23		2	3		
24		2	2		
25		2	3		
26		2	1		
27		2	3		
28		2	1		
29		2	2		
30		2	3		
31		2	2		
32		2	3		
33					

「データ分析」の選択

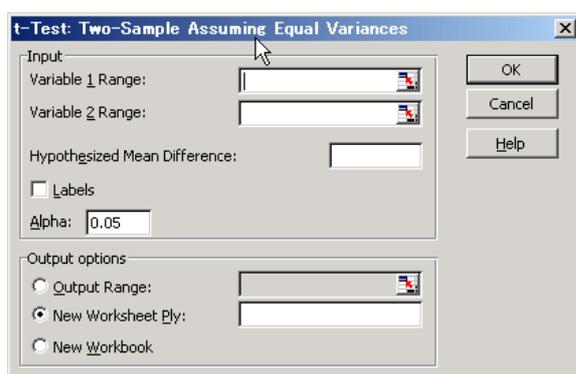
- データ > データ分析 を選択する。



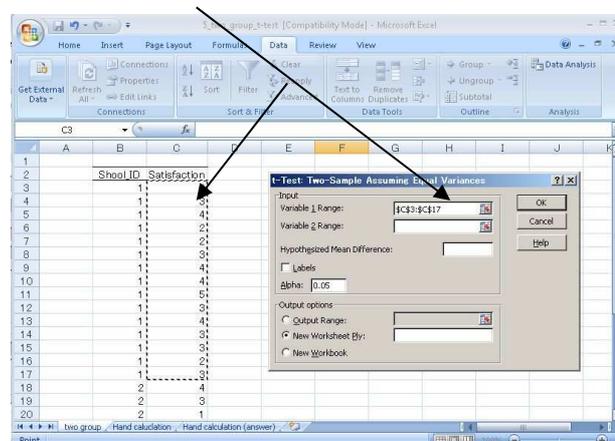
- 「等分散を仮定した2サンプルのt検定」を選択する(t-Test: Two-sample Assuming Equal Variance)⁸。



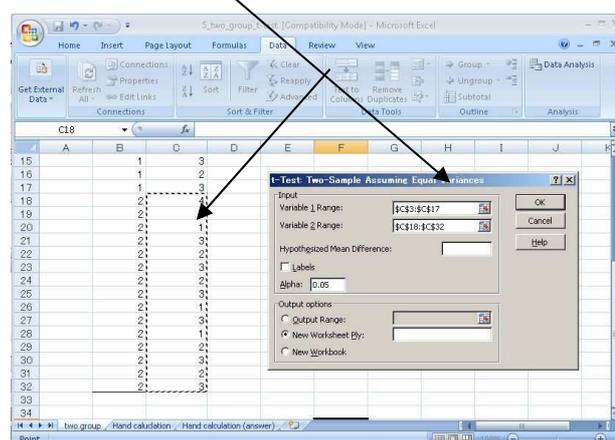
- 以下のボックスが現れる。



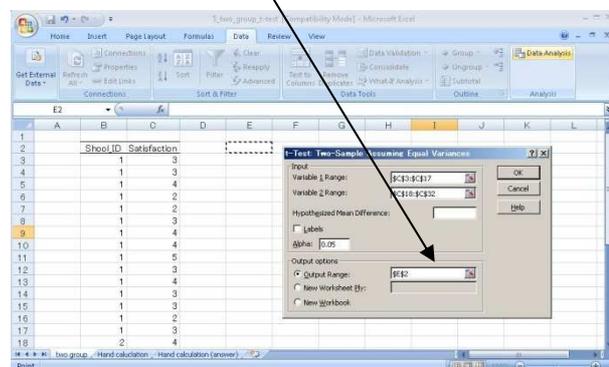
- 我校(School_ID 1)のデータを選ぶ。



- 近隣校(School_ID 2)のデータを選ぶ。



- 出力オプション>出力範囲を選択して、空いている適当なセルを選ぶ。



⁸ 等分散の F 検定により等分散ではないと判断された場合には、「異分散を仮定した2サンプルの t 検定」を選択する(t-Test: Two-sample Assuming Unequal Variance)。

分析結果を得る。

Shool_ID	Satisfaction
1	3
1	3
1	4
1	2
1	2
1	3
1	4
1	4
1	5
1	3
1	4
1	3
1	3
1	2

	Variable 1	Variable 2
Mean	3.2	2.4
Variance	0.7429	0.8286
Observations	15	15
Pooled Variance	0.7857	
Hypothesized Mean	0	
df	28	
t Stat	2.4717	
P(T<=t) one-tail	0.0099	
t Critical one-tail	1.7011	
P(T<=t) two-tail	0.0198	
t Critical two-tail	2.0484	

t-値
 $t > 2$: 有意
 $t \leq 2$: 有意とは言えない

p-値 (確率値, %)
 $p < 0.05$: 有意
 $t \geq 0.05$: 有意とは言えない

結論

“適用グループ”（我校）の平均値と“比較グループ”（近隣校）の平均値の差は統計学的に有意である（ $p < 0.05$ (5%)）。したがって、事業の実施によって、状況は改善したと判断できる。

【発展学習：2群の有意差検定の手計算】

Independent t-test (Two-groups comparison)

Town_ID	Satisfaction	$(X_{1i} - \bar{X}_1)$	$(X_{1i} - \bar{X}_1)^2$
1	3	-0.2	0.04
1	3	-0.2	0.04
1	4	0.8	0.64
1	2	-1.2	1.44
1	2	-1.2	1.44
1	3	-0.2	0.04
1	4	0.8	0.64
1	4	0.8	0.64
1	5	1.8	3.24
1	3	-0.2	0.04
1	4	0.8	0.64
1	3	-0.2	0.04
1	3	-0.2	0.04
1	2	-1.2	1.44
1	3	-0.2	0.04

Town_ID	Satisfaction	$(X_{2i} - \bar{X}_2)$	$(X_{2i} - \bar{X}_2)^2$
2	4	1.6	2.56
2	3	0.6	0.36
2	1	-1.4	1.96
2	3	0.6	0.36
2	2	-0.4	0.16
2	3	0.6	0.36
2	2	-0.4	0.16
2	3	0.6	0.36
2	2	-0.4	0.16
2	1	-1.4	1.96
2	3	0.6	0.36
2	1	-1.4	1.96
2	2	-0.4	0.16
2	3	0.6	0.36
2	2	-0.4	0.16
2	3	0.6	0.36

Total = 48

Total₁ = 10.4

Total = 36

Total₂ = 11.6

Average $\bar{X}_1 = 3.2$

Average $\bar{X}_2 = 2.4$

$n_1 = 15$

$n_2 = 15$

$$s_p^2 = \frac{\text{Total}_1 + \text{Total}_2}{(n_1) + (n_2) - 2} = 0.7857$$

$$s_p = \sqrt{s_p^2} = 0.886$$

$$SE = s_p \cdot \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = 0.3237$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{SE} = 2.4717$$

$p\text{-value} = 0.0198$

* p -value is obtained by EXCEL Command: =tdist (t, n1+n2-2, 2)
In this case, type [=tdist(2.471658,28,2)] at any cell. You can get 0.019795.

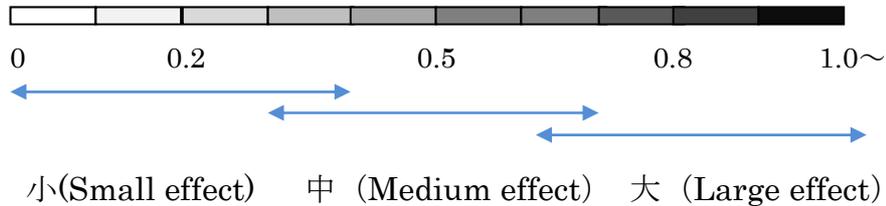
<Advance: Effect size (Cohen's d)>

$$d = \frac{|\bar{X}_1 - \bar{X}_2|}{S_p} = 0.90252$$

Criteria: Small: $d = 0.2$, Medium: $d = 0.5$, Large: $d = 0.8$ (Source) Cohen (1988)

別添4：エフェクトサイズの計算

さらに、エフェクトサイズ（効果サイズ）（ d ）を計算して、改善効果の大きさを判断することは有益である。矢印が重なっている部分はグレーゾーンであり、中間的な書き方（例：「小～中」「中～大」とすることが勧められる。



(出所) Cohen, J. (1988)をもとに筆者作成

算出した結果、エフェクトサイズ（効果サイズ）（ d ）が0～0.4であれば、事業の効果は「小」、0.3～0.7であれば、事業の効果は「中」、0.6以上であれば、事業の効果は「大」と判断される。

エフェクトサイズ(d)の計算式は次のとおり。

$$d = \frac{|\bar{X}_1 - \bar{X}_2|}{S_p} = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\text{Pooled Variance}}}$$

この例だと・・・

↓エクセルの出力結果を参照する

	Variable.1	Variable.2
Mean	3.2	2.4
Variance	0.742857143	0.828571429
Observations	15	15
Pooled Variance	0.785714286	

$$d = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\text{Pooled Variance}}} = \frac{3.2 - 2.4}{\sqrt{0.7857}} = 0.9025$$

したがって、改善効果は「大」と判断できる。

執筆者たちの紹介

<p>著者</p> <p>佐々木亮</p> <p>国際開発センター (IDCJ) 評価部主任研究員</p>		<p>NGO セミナー企画・運営 & ガイドライン編集担当</p> <p>小室雪野</p> <p>国際開発ソリューションズ ／日本評価学会研修委員</p>	
<p>最終学歴： ウェスタンミシガン大学評価研究所 評価学博士 ニューヨーク大学大学院公共行政学修士</p> <p>専門分野： 評価論理、統計分析、NGO 評価</p> <p>論文・出版物： 「評価論理：評価学の基礎」(2010)、「エクセルで政策評価」(2007)、「政策評価トレーニングブック」(2003)、「政策評価の理論と技法」(2000, 2004 増補改訂版)、その他の調査報告書多数</p> <p>Dr. Ryo SASAKI</p> <p>< Academic Background ></p> <ul style="list-style-type: none"> - Ph.D. in Evaluation, the Evaluation Center, Western Michigan University - Master of Public Administration (Public Policy Analysis), New York University <p>< Current position ></p> <ul style="list-style-type: none"> - Senior researcher, IDCJ 	<p>最終学歴： 東京 YMCA 英語専門学校 法政大学経済学部（通信教育課程）中退</p> <p>専門分野： ODA 評価、評価研修の企画・運営</p> <p>論文・出版物： 「日本評価学会「評価士養成講座」の実施：評価の専門能力の認証制度の確立に向けて」(2012、共著)、その他の調査報告書多数</p>		