

Intervention Studies

ぎっくり腰になったら休みますか？それとも働きますか？

intervention studies（介入研究）は原因と結果の証拠を直接引き出す為に行なわれます。Study population を絞り込むことによって、より明確な結果を得ることができるかもしれませんが、generalizability を失うことにもなります。Study population の設定はその辺の trade off なのです。

私は腰痛もちで不用意に自転車を持ち上げたときなど“グキッ”となって何度かぎっくり腰になったことがあります。職場から仕事にいかないわけにもいかず、非アスピリン系の薬をごっそりと飲んで仕事を続けます。このつらさは腰痛持ちにしか判りません。ある人は「腰はこわいから休んで寝ていたほうがいいんじゃない？」とやさしい言葉をかけてくれるかもしれませんが、「元気そうだな。腰が痛いくらいで休むなよ。」ときびしい言葉がかえってくるかもしれません。医者の不養生という言葉がありますが、本当はどうするべきなのでしょう？普段よく経験する病態ですが、医者によっての治療方針はばらばらで、本当のことは以外に知られていません。

The Treatment of Acute Low Back Pain – Bed Rest, Exercises, or Ordinary Activity?

Malmivaara A. et al. New Eng J Med. 332:351-5,1995.

腰痛治療 - ベッド上安静、運動、日常生活？

要旨

ベッド上安静と背筋を鍛える運動が腰痛の際推奨されます。この相反する治療の有効性については不明のままです。そこで我々は腰痛を主訴に職員ヘルスケアセンターを受診したヘルシンキの市職員を対象に臨床試験を行ないました。患者さんはランダムに3つの治療、すなわち2日間のベッド上安静(67人)、背筋を使う運動(52人)、耐えられる程度の日常業務継続(52人)に振り分けました。3と12週後に調査をしたところ、日常業務を継続した患者さんにおいて、腰痛を感じる期間、程度、腰部伸展の程度、仕事に支障を来たしたか、欠勤の日数において有意に良好な結果を得ました。回復はベッド上安静が最も悪い事が判明しました。腰痛患者さんに対しては、通常業務継続が他の治療より早期回復を促しました。

質問

1. 研究対象は？場所、時間、個々の特徴は何ですか？
2. 何故諸々の患者さんが除外されたのですか？
3. このような研究(randomization)においてどの程度まで confounder を考慮しなくてはなりませんか？
4. 比較群間の予後因子を比較した際、baseline に差を認めたとすると、またこれらの偏りが単に確率(chance)で発生したとすると、結果を歪めることになりませんか？
5. 全ての患者の特徴が調査された後、治療内容からして、患者と治療者はどのような治療を受けているか知っています。著者らはどうしてこのような手段を取ったのですか？
6. 治療のルールを守らない事によって、治療効果測定結果にどのような影響を与えますか？
7. 結果の測定方法に risk や rate を使用していませんが、何を結果測定に用いているのですか？治療グループは cohort と言えますか？
8. Table 3 にある欠勤の日数を考えた時、補正されたグループの平均の相違は補正していないデータの相違とどのような点で異なりますか？このことは randomization した後も confounding が残存しうることになるのですか？

緒言

腰痛はしばしばある訴えで、医師は往々にしてベッド上安静を指示しますが、その真偽は判っていません。そこで我々は randomization による臨床試験を行ないました。

方法

腰痛を主訴に職員ヘルスケアセンターを受診したヘルシンキの市職員を対象に臨床試験を行ないました。急な腰痛あるいは3週間以上継続する腰痛を訴えて受診した場合に対象に含め、60度以下でラセーグ兆候陽性あるいは神経症状を呈し椎間板ヘルニアを疑わせる患者さん、妊婦、癌、骨折、泌尿器疾患患者さんは除外してあります。このような条件に合った患者さんに対して、看護婦が臨床試験の目的と方法を記載したものを渡し、この3つの治療の有効性の優劣に関しては不明である点を説明し、試験に参加するか否か本人に意思決定してもらっています。

腰痛の患者さんで、試験参加の意思表示をした時点で、患者さんは3つのうちのどれかの治療が書かれた紙の入った封をしてある封筒を受け取ります。そしてヘルスケアセンターの医師が診察を行ない試験に適合するかどうかを審査します。その後その医師は封を開け、治療法を説明します。その後経過観察(次回受診)の必要が無い程度の軽症のものも研究対象から除外しました。(注* randomization の際は randomization を行なう前にどの治療に割り当てられるか判らないことを説明した上で参加の意思決定をしてもらうと同時に、健康状態などをチェックして参加適応があるかの判断も行なわなくてはなりません。もしも randomization の後副作用等以外の理由で対象からはずれること、あるいははずすことのないようにしなくてはなりません。)

ベッド上安静では、2日間の絶対安静で、トイレ歩行等のみ可としました。患者さんには安静の際の semi-Fowler position をイラストで説明しています。2日間の安静のあとは日常生活に戻ってもらいます。運動群にあたった患者さんは理学療法士より背筋運動の方法の説明を受けています。日中毎時間背筋運動を痛みが出現するまで行なうものです。背筋運動は前横後ろなど各方向に10回ずつゆっくりと動かすよう指導しています。日常生活群では、腰痛の程度が許すかぎりなるべく普段の生活をしてもらうように頼んであります。運動群と日常生活群ではベッド上安静をなるべく避けるように指導してあります。

本臨床試験の方法に関してはフィンランドの Institute of Occupational Health Patients における倫理委員会で承認されています。(注* intervention study では倫理面が問題で、その点どのように対処したかを書き示す必要があります。)

質問用紙にて本臨床試験のプロトコールが正しく遂行されたか確認しています。3週間後と12週間後に医師を再診し、腰痛を含む状態をチェックしてもらいます。再診しなかった患者さんに対しては電話連絡をとり、再診を要請します(注*なるべく missing data を減らす努力をしなくてはなりません)。結果の評価は、質問用紙、理学療法士の判定、カルテによってなされました。また結果を判断する人間に、本研究の結果が判る前に、3つの治療法のどれが有効であると思うかについても質問してあります。

パワーを0.80、 $\alpha = 0.05$ に設定すると各グループで64症例が必要とされました。3つの治療効果の判定には analysis of covariance (ANOVA) を用いました。

結果

全部で186人が本臨床試験に参加してくれました。3週間後の追跡率は89%であり、

ベッド上安静群では 5 人、運動群では 10 人、日常生活群では 6 人の欠席がありました。12 週後の追跡率は 87% で、それぞれ 8 人、11 人、5 人が欠席しました。3 週目、12 週目に受診した人としなかった人の特徴の間に特別な差を認めませんでした。16 人は 186 人から除外されました。その理由は、質問用紙未提出であった、あるいは医師が適合基準に合致すると判断したが後に合致しないことが判明したことが理由です。これらの判定は、治療内容を知らない状態で除外を判断しています。16 人中 9 人が運動群、4 人がベッド上安静、3 人が日常業務群でした。

Table 1

	ベッド上安静 (N=65)	運動 (N=52)	日常生活 (N=67)
年齢	40.8	41.1	39.1
女性(%)	60	71	70
教育(年)	11.4	11.3	11.2
結婚(%)	39	48	46
Body Mass Index	25.0	25.2	25.0
Weight(kg)/[height(m)] ²			
運動(週 3 回以上) %	54	54	58
月収 \$	1858	1947	1852
1 日 5 時間以上の重労働%	10	13	22
仕事の満足度	18	23	25
痛みの程度と仕事の支障度			
痛みの持続(日)	4.7	5.1	4.5
痛みの強さ(score 0 – 10)	5.9	6.1	5.7
膝より下までの放散痛%	21	10	9
欠勤日数	1.4	1.1	1.4
仕事能率(score 0 – 10)	4.9	5.3	4.8
足上げの角度	73	74	72
腰曲げ(cm)	5.1	4.9	5.1
Oswestry back-disability index	34.6	33.8	32.0
Health-related quality of life index	0.86	0.86	0.85
腰痛の既往(過去 12 ヶ月)			
30 日以上の腰痛%	16	27	15
7 日以上の腰痛%	14	13	11
医療機関受診回数	0.82	0.64	0.55
腰部手術の既往%	0	2	0

日常生活群は重労働が若干多く、ベッド上安静群では膝より下までの放散痛が若干多く、運動群では 30 日以上の腰痛が若干多い傾向にありました。運動群の 2 人で腰部手術の既往がありました。3 群全ての患者さんの職種は多岐にわたっていました。

Compliance(どれくらい指示に従えたか) 注* もちろん自己申告だと正しいとは限りません。)

3 週間の追跡期間、日常生活群が 2 時間のベッド上安静(昼間)であったのに対して、ベッド上安静の患者さんは 1 日平均 22 時間でした。日常生活群が決められた運動を 1 日平均 3 セット行なっていたのに対して、運動群は 1 日平均 61 セットこなしていました。ベッド上安静、運動、日常生活群それぞれに対しての鎮痛剤処方率は 93%, 91%, 93%

でした。

腰痛を判定したヘルスケア従事者の治療法に対する意識 (注*今回はdouble blind study ではないので治療者側に著しい意識の偏りがあるかないかを知ることは参考になります。)

治療が始まる前に、医師、看護婦、理学療法士の3つの治療法についての考えをきいたところ、36人中10人が「運動療法が一番有効」と答え、3人が「ベッド上安静」、3人が「日常生活」を答え、残りは特に治療に大きな差を持たないだろうと予想していました。

3週間後の結果 (注*参加者の主観的症状のみでなく欠勤日数や医療機関受診回数は客観的事実としてとらえることができます。)

基本的な特徴(base-line measurement; Table 1)の違いを補正したところ、日常生活群は仕事能率と欠勤日数の面で有意にベッド上安静群より優れていました。一方、日常生活群は、欠勤日数、腰痛持続期間、Oswestry index の面で運動群より優れていました。欠勤日数はベッド上安静、運動群で中央値 5 日であり、日常生活群では中央値 4 日でした。しかし1週間後欠勤状況は、ベッド上安静で41%、運動群で36%、日常生活群で20%でした。さらに2週間後の欠勤状況はそれぞれ19%、11%、2%、3週間後のそれは5%、6%、2% でした。Fisher's exact test (次項参)_を行なったところ、ベッド上安静群と日常生活群の間で1週間後、2週間後にそれぞれ $p=0.01$, $p=0.02$ の有意な差を認めました。

*** Fisher's exact test ***

標本数が少ない時、正規分布がはっきりしないことが多いので、通常の方法では確認できません。

仮に以下のようなデータを得たとしましょう。

死亡原因	ベッド上安静	日常生活	合計
欠勤	19 (a)	2 (b)	21 (a + b)
出勤	43 (c)	59 (d)	102 (c + d)
合計	62 (a + c)	61 (b + d)	123 (a + b + c + d = n)

$$\Pr (a, b, c, d) = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!}$$

$$\Pr (19,2,43,59) = \frac{21! 113! 62! 61!}{123! 19! 2! 43! 59!} < 0.002$$

よって「ベッド上安静の人が欠勤する確率と日常生活をした人が欠勤する確率は同じである」とする H0 は棄却され、上 2 者は異なると結論できます。

Table 2

	ベッ ド上 安静 (N=6 2)	運動 (N=4 2)	日常 生活 (N=61)	ベッド上 - 日常生活	調整後 運動 日常生活
痛みの程度と仕事の支障度					
痛みの持続(日)	7.5	5.7	4.1	3.2 (1.3-5.0)	1.8 (0.1-3.5)
痛みの強さ(score 0 – 10)	15	20	14	0.3 (-3.1-3.8)	4.3 (0.8-7.8)
膝より下までの放散痛%	2.4	3.1	1.9	0.3 (-0.4-0.9)	0.9 (-0.01-1.7)
欠勤日数	13	10	8	-1.5 (-29.6-26.6)	-2.0 (-33.2-29.1)
仕事能率(score 0 – 10)	6.8	7.2	7.9	-0.9 (-1.8-0.02)	-0.3 (-1.2-0.5)
足上げの角度	86	86	87	0.3 (-4.2-4.7)	-0.03 (-5.1-5.0)
腰曲げ(cm)	6.2	6.0	6.4	-0.3 (-0.7-0.2)	-0.3 (-0.7-0.1)
Oswestry back-disability index	16.0	18.6	10.0	3.9 (-0.2-8.0)	6.6 (2.0-11.1)
Health-related quality of life index	0.92	0.91	0.94	-0.02 (-0.04-0.002)	-0.02 (-0.04-0.001)
治療満足度	7.0	7.7	7.6	-0.7 (-1.8-0.4)	0.5 (-0.6-1.6)
Compliance					
1日のベッド上安静の時間(H)	22	5	2		
運動回数(セット)	8	61	3		

12 週後の結果

Table 3

	ベッ ド上 安静 (N=5 9)	運動 (N=4 1)	日常 生活 (N=6 2)	ベッド上 - 日常生活	調整後 運動 日常生活
痛みの程度と仕事の支障 度					
欠勤日数	9.2	7.2	4.7	3.4 (0.2-6.5)	2.5 (0.2-4.9)
痛みの強さ(score 0 - 10)	2.1	1.8	1.3	0.7 (0.03-1.4)	0.2 (-0.5-1.0)
膝より下までの放散 痛%	14	10	11	4.1 (-13.4-21.6)	-0.01 (-28.2-26.2)
仕事能率(score 0 - 10)	7.7	7.8	8.5	-0.8 (-1.5-0.1)	-0.7 (-5.8-4.4)
足上げの角度	88	89	90	0.4 (-4.8-5.5)	-0.7 (-5.8-4.4)
腰曲げ(cm)	6.3	6.0	6.6	-0.6 (-1.1-0.1)	-0.6 (-1.1-0.1)
Oswestry back-disability index	11.8	10.8	7.4	3.8 (0.1 - 7.5)	2.6 (-1.6 - 6.7)
Health-related quality of life index	0.93	0.95	0.95	-0.02 (-0.04-0.002)	-0.02 (-0.04-0.001)
治療満足度	7.3	8.1	7.7	-0.6 (-1.6-0.4)	0.4 (-0.6-1.4)

Study Questions

1. 研究対象は？場所、時間、個々の特徴は？

1992 年から 1993 年 4 月までの間、ヘルシンキ在住市民の中から腰痛を訴えて職業健康センターを受診した市職員を対象としました。椎間板ヘルニア、妊娠、癌、骨折、泌尿器疾患を持つ腰痛患者を除外しています。また次回受診の必要のないものも除外。交通機関、電気関係の職員を除外した理由は不明です。おそらくこの人達は自社の診療所を持ち、職業健康センターを受診しなかったのではないかと想像します。

2. 何故諸々の患者さんが除外されたのか？

上記疾患による腰痛は予定されている治療に対して抵抗性を示す可能性が高く、有意差を検出できないかもしれません。また少なくとも以下の3つのことがいえます。

- i. 筆者らが調査しようとしているのは一般的な腰痛であり、特殊な疾患に伴う腰痛ではありません。
- ii. 治療に反応しない腰痛が含まれると治療効果の程度および方向性を正しく評価できなくなってしまう。
- iii. 治療に反応しない腰痛が含まれると、治療群とコントロール群で imbalance を生じた場合、治療効果の評価を歪めるかもしれません。

最初条件に合った腰痛患者として臨床試験に含めて経過観察されていた患者さんのうち、後で問題ありとして除外されたものがありました。除外は randomization の前に行なわれるべきです。

3. このような研究においてどの程度まで confounder を考慮しなくてはならないか？
我々は Randomization によって confounder が除外されることを期待します。しかし本臨床試験のように数が十分に多くない場合、治療群の間で base line characteristics に差を生じる事があります。よって randomization をしたからといって confounder を完全に除外することはできません。

通常は実数と SD を示します。小さな研究における有意差のない違いは結果を歪める可能性がありますし、逆に大きな研究における有意差のある小さな違いは結果を歪めるとはかぎりません。よって confounding の程度は必ずしも有意差の存在に関係しないので、表 1 の各群の人数の偏りにおいて、p 値を示す必要はありません。

4. 比較群間の予後因子を比較した際、baseline に差を認めたとすると、またこれらの偏りが単に確率(chance)で発生したとすると、結果を歪めることになりませんか？
Confounder として結果を歪めます。よって stratification や statistical modeling によって補正が必要となります（臨床試験の統計学を参照）。

5. 全ての患者の特徴が調査された後に患者と治療者は治療内容を知る仕組みとなっている。何故この臨床試験を企画した者達はこのような方法を選んだのか？
Randomization の後に発生する全ての事象は、どの治療に割り当てられたかに影響されると考えられます。それは治療が始まる前においても同様です。よって baseline 測定は randomization の前に行なわれなくてはなりません。

6. 治療のルールを守らない事によって、治療効果測定結果にどのような影響を与えるか？
日常生活群と他の 2 つの群の差が縮まるため、治療効果の差の程度が小さくなります。一方ベッド上安静群（あるいは運動群）が日常生活を送ったら、運動群（あるいはベッド上安静群）より腰痛の治りがよいことになってしまいます。2 群間の Non-differential misclassification においては null hypothesis の方向に向かいますが、3 群間では必ずしもそうとは限りません。Differential misclassification でもどちらの方向に歪むかわかりません（差が無くなる場合もあれば、むしろ差が広がることもあるということ）。

7. 結果の測定方法に risk や rate を使用していないが、何を結果測定に用いているのか？治療グループは cohort と言えるのか？
臨床試験を含めた cohort study では risk や rate を結果測定に用いる事が多いのですが、これらは cohort study の定義にはなりません。本臨床研究では 3 週と 12 週の（累積）実測値をみています。

8. 表 3 にある sick days の日数を考えた時、補正されたグループの平均の相違は補正していないデータの相違とどのような点で異なるのか？このことは randomization した後も confounding が残存しうることになるのか？
ベッド上安静と日常生活の欠勤日数の差は 4.5 日ですが、baseline の違いを調整した

あとは 3.4 日でした。このことはベッド上安静群に、欠勤が長引きやすい条件（例えば腰痛の予後不良因子）を持つ患者さんが多目に分布していたことを示唆しています。もし baseline の相違を補正したならば、その相違による影響は減少します（除去されるとはいいきれません）。

コメント

Randomized study であっても confounder が発生しうることを知ってください。