

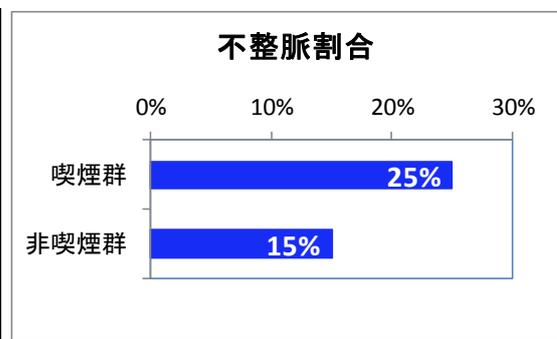
リスク比・オッズ比の解説書

統計分析研究所
株式会社アイスタット代表取締役社長
ビジネス・ブレークスルー大学大学院授
(理学博士) 菅民郎

医師: 志賀くん、いつもわかりやすく情報提供してくれているのでとても助かっているよ。ありがとう。ところで、すごくいいタイミングで来てくれたんだけど、ちょっとこのデータについて、相談にのってくれないかな…。

調査データ

	不整脈		横計	割合
	ある	ない		
喫煙群	48 人	143 人	191 人	25.1%
非喫煙群	39 人	220 人	259 人	15.1%
縦計	87 人	363 人	450 人	19.3%



志賀: えっ、あっ、はい(どきどき)

医師: タバコを吸っている患者さんには、とにかく禁煙をすすめているし、循環器疾患の観点からも、指導するようにしているのだけど、まだまだタバコをやめることができない患者さんもいてね。ボクももう1回、喫煙がおよぼす疾患への影響について再度勉強しようと思って、患者さんのカルテと問診からこんな数表をつくってみたんだよ。

志賀: うわあー先生、こんなに多くの患者さんのデータをですか！ 毎日の診療だけでも忙しいのに、大変でしたでしょうね。

医師: いやいや、そんなことはないよ。なによりも病気になってからじゃおそいからね。少しでも患者さんが良くなってくれればと思っているんだよ。

志賀: 先生、すごいなあ(感心)。

医師: けっこう頑張って集めて整理したんだよ。今までもこのようなデータはみていたし、簡単にわかったつもりでいたのだけど、よくよくみると、ごちゃごちゃになってしまったんだよ。志賀くん、ぜひ整理して教えてもらえないかな。

志賀: え〜と、どこでしょうか？

医師: このような数表と横棒グラフをつくってみたのだけど、ここから、喫煙と不整脈の有無とのリスク比やオッズ比とか、有意差の出し方がよくわからなくなってしまっ…。恥ずかしい話

なんだけど、それぞれどういうことなのか、わかりやすく教えてもらえないかな？

最近、疫学試験などでもよく目にするもののあるリスク比とオッズ比。いつもながらもっとデータについて勉強しなければ、と思った志賀さん。「先生、きちんと調べて、ご説明に伺います」と話し、医師も「また、わかりやすく、お願いね！」と期待してくれた様子であった。そこで志賀さんは、先日来教を求めているビジネス・ブレイクスルー大学院教授で統計学の専門家・菅民郎先生を訪ねた。

志賀さんは菅先生に、循環器専門医からの宿題を説明。菅先生はニコニコ笑いながら、話を聞いてくれました。そしてリスク比とオッズ比、有意差の出し方などについて、いつものようにとてもわかりやすくアドバイスしてくれたのでした。



ポイント1: 分割表とは何だろう

志賀: 菅先生、いつも大変お世話になっております。またいろいろとお教えいただきたいことがあってお邪魔させていただいたのですが…。

菅: 志賀さんはいつも仕事熱心だねえ。どれどれ。

志賀: このデータからリスク比とオッズ比、有意差の出し方についてお教えいただきたいのです。

菅: なるほど、よしじゃあ、最初に簡単な例で説明しよう。

志賀: “簡単”という言葉大好きです。

菅: いつものように難しく考えないで、「簡単に」でいこう！ 次の表は、10人の患者について、「不整脈があるかないか」「喫煙しているか、していないか」を調べたものとして。

データは、「喫煙」を1、「非喫煙」を0、不整脈が「ある」を1、「ない」を0としているよ。

表 1

患者No.	喫煙有無	不整脈有無
1	0	0
2	1	1
3	0	1
4	0	0
5	1	0
6	0	0
7	0	0
8	1	0
9	1	1
10	1	1
	1 喫煙 0 非喫煙	1 ある 0 ない

菅: 志賀さん。この表から、不整脈について喫煙者と非喫煙者を比較したとき、両者に差があるかどうかを明らかにできるかな。

志賀: この表をながめていても傾向が分からないので、まずはデータを並べ替えて見るといいんじゃないでしょうか？

菅: ほう、どんなふうに並べ替えるのかな。

志賀: 喫煙有無別、不整脈有無別に並べ替えてみます。

表 2

患者No.	喫煙有無	不整脈有無
2	1	1
9	1	1
10	1	1
5	1	0
8	1	0
3	0	1
1	0	0
4	0	0
6	0	0
7	0	0
	1 喫煙 0 非喫煙	1 ある 0 ない

菅: 志賀さん、いいねー。だけどこの表から何が分かるのかな。

志賀: 喫煙者が5人いて、そのうち3人は不整脈がある人です。非喫煙者は5人いて、そのうち1人は不整脈がある人です。

菅: ということは？

志賀: えーと……。

菅: 喫煙者の有無別に、不整脈のある人の割合を計算してみたらどうなる？

志賀: はい、喫煙者における不整脈のある人の割合は3/5で60%です。非喫煙者における不整脈のある人の割合は1/5で20%です。ですので、不整脈のある人の割合は、喫煙者が60%、非喫煙者が20%で、喫煙者の方が40%高いです。このため、「喫煙者と非喫煙者を比較したとき、不整脈において差があるといえる」と思います。

菅: 志賀さん、いいねー。ここまでは合格だよ。ただ、「差がある」かどうかは、有意差検定をする必要があるんだけど、それはこれからの話だ。

志賀: 以前教えてもらったp値ですね。

菅: それもあるけど、今日はリスク比とオッズ比ということだったね。

志賀: あっ、そうでした！



ポイント2: 因果関係のクロス表 表の左側に原因、上側に結果を書く！

菅: では、志賀さんが集計した結果を表にしてみてくださいかな。

志賀: はい。(シコシコ表にしてみる)これでよいでしょうか。

表3 分割表

	不整脈有無		横計	割合
	ある	ない		
喫煙	3	2	5	60%
非喫煙	1	4	5	20%

菅: OK。この表のことを分割表(contingency table)というんだ。

志賀: この表(表3)は、表の行(左側)に喫煙有無、列(上側)に不整脈有無としたのですが、次の表(表4)のように行と列を入れ替えて、表の行(左側)に不整脈有無、列(上側)に喫煙有無にした方がいいですか？

表4 行と列を入れ替えた分割表

		喫煙	非喫煙
不整脈有無	ある	3	1
	ない	2	4
縦計		5	5
割合		60%	20%

菅: それはだめなんだ。

志賀: えっ、そうなんですか？！

菅: 因果関係を考える場合、原因と結果があるでしょ。原因と結果の関係を調べるために分割表を作る場合、行(左側)に原因、列(上側)に結果の項目を置くのがルールなんだ。そして、「ある」を横計で割り、割合を求めるんだ。

志賀: へー、偶然に正しい表(表3)を作ってしまった。(笑)



ポイント3:リスク比とは何だろう

菅:そしてしっかり理解してほしいのは、この表3の「ある」を横計で割って得られた「割合」のことをリスク(Risk)と言うんだ。

志賀:リスクといえば、「危険」や「恐れ」という意味ですよね…。

菅:そうだね。

志賀:えーと、つまり、今回のケースでの“リスク”は、不整脈になる“危険”や“恐れ”が喫煙によってどの程度あるのかがわかる、ということでしょうか。

菅:志賀さんはすごいなー、そうだよ。では、リスクを喫煙者、非喫煙者で計算してみようか。

志賀:はい。(表3をじっくりながめながら…)喫煙者が不整脈となるリスクは $3 \div 5$ で60%、同様に非喫煙者が不整脈となるリスクは20%、ということです。

菅:正解! リスクの差はどうなる?

志賀:えーと、 $60\% - 20\%$ で40%なので、「リスクは喫煙者が非喫煙者を40%上回っている」で良いでしょうか。

菅:完璧だよ。

志賀:それほどでは(やったあー)

菅:じゃあ次に、喫煙者のリスクを非喫煙者のリスクで割ってみてくれるかな。

志賀: $60\% \div 20\%$ だから、3です。

菅:そう、その値がリスク比(Risk Ratio)なんだ。

志賀:リスク比って、こんなに簡単に求められるんですか。

菅:そうだよ。だから大事なのは求め方ではなく、解釈の仕方なんだよ。

志賀:はい。

菅:リスク比3から次のことがいえるんだ。「喫煙者が不整脈となるリスク(割合)は非喫煙者に比べ3倍である」

志賀:なるほど、リスク比はゲットしました!!



ポイント4:オッズ比とは何だろう。

菅:よし、オッズ比を勉強してみよう。オッズは競馬とかギャンブルでよく使われるから、勝負事の例で話をしようかな。そのための材料としたいんだけど、志賀さんの持っているスマホにあるアプリゲームをなんでもいいから、5回やってみてくれるかな。

志賀:遊んでしまっていていいんですか。

菅:遊びながら勉強できるんだから、いいんじゃないか、でも1ゲーム1分とかからないのにならうか。志賀さんの後に、私もやってみよう。そしてどちらが強いかを確かめられないか。

下記の表は二人の成績を示したものである。

表5 成績表

	アプリ/ゲーム		ゲーム数
	勝数	負数	
志賀さん	3	2	5
菅先生	1	4	5

志賀:先生、偶然にも喫煙有無と不整脈有無との関係と同じ結果になりましたね。

菅:そうだね。よし、オッズの説明をする前にリスク比のおさらいをもう一度しておこう。志賀さん、この表についてリスク比を求めて、リスク比から何がいえるかを考えてみてくれるかな。

志賀:はい(カリカリカリ)。できました。

表6 成績表(リスク比あり)

	アプリ/ゲーム		ゲーム数	勝率
	勝数	負数		
志賀さん	3	2	5	60%
菅先生	1	4	5	20%

志賀:私の勝率は60%です。菅先生の勝率は20%です。勝率の比、すなわちリスク比は、 $60\% \div 20\%$ より3となる。リスク比から、私の勝率は菅先生に比べ3倍である。勝率をゲームの強さと考えると、私のゲームの強さは菅先生より3倍だといえます。

菅:よくできました。志賀さん勝負強いね。それでは、この成績表を用いてオッズ比について説明しようか。

志賀:お願いします。

菅:志賀さんの勝数を私の勝数で割った値をオッズ(Odds)というんだ。そして、志賀さんの負数を私の負数で割った値もオッズというんだ。

志賀:ええっ、ちょっと待ってください。どっちもオッズっていうんですか。

菅: あせらない、あせらない、もうちょっと続きの説明を聞いて。勝数に着目すると、志賀さんの勝数(3勝)は私(1勝)に比べ3倍、すなわち、勝数オッズは3である。負数に着目すると、志賀さんの負数(2敗)は私(4敗)に比べ半分、すなわち、負数オッズは0.5である。

志賀: はい。ここまではわかります。

菅: そして、勝数オッズと負数オッズの比をオッズ比(Odds Ratio)という。

志賀: なるほど、勝数オッズと負数オッズっていうんですね。ということは、オッズ比は $3 \div 0.5$ より、6ですね。

菅: そのとおり。

志賀: オッズ比から、私の勝率が菅先生に比べ6倍、私のゲームの強さは菅先生より6倍だということですね！

菅: ブー。絶対だめだ。大間違い！ここが、多くのひとが陥るオッズ比の典型的な間違った解釈なんだよ。勝率(強さ)の比較はリスク比でしかできないんだ。

志賀: ええ——、では、オッズ比から何が分かるんですか。

菅: オッズ比の値が大きいとか小さいといったことが分かるだけなんだよ。オッズ比はリスク比に比べ理解しにくく、だから使い方に注意が必要なんだ。

志賀: そうなんですか。

菅: ここで、今までのことが理解できたかをテストしてみよう。

志賀: ヒェー、そんなあ。(汗)

菅:では、喫煙有無と不整脈有無の簡単事例の分割表(表3)があったよね。これについてリスク比とオッズ比を求めて、解釈してくれるかな。

志賀:はい。(カリカリカリ。そして以下の表にまとめました)

表6 喫煙状況と不整脈のリスク比とオッズ比

	不整脈		横計	リスク
	ある	ない		
喫煙	3	2	5	60%
非喫煙	1	4	5	20%
オッズ	3.0	0.5		

リスク比	3
オッズ比	6

志賀:喫煙者の不整脈のリスクが60%、非喫煙者の不整脈のリスクは20%ですので、リスク比は $60\% \div 20\%$ から3です。ですので、喫煙者が不整脈となるリスクは非喫煙者に比べ3倍であるといえます。

菅:そうだね。

志賀:オッズ比にまいります。不整脈があるケースのオッズは、喫煙者が3、非喫煙者が1のため、 $3 \div 1$ からオッズは3。不整脈がないケースのオッズは、喫煙者が2、非喫煙者が4のため、 $2 \div 4$ からオッズは0.5。これより、オッズ比は $3 \div 0.5$ で6となりました。オッズ比の値は6と大きいので喫煙有無は不整脈の影響要因といえそうですが、喫煙者が不整脈となるリスクは非喫煙者に比べ6倍だと言ってはいけない、ということでしょうか。

菅:大変よくできました。

志賀:うーん、先生、オッズ比はリスク比と比べるとあまり使い道がないように思うんですけど…。でも臨床研究の論文では、オッズ比はよく使われています。

菅:理解しにくいオッズ比がなぜ、臨床研究で使われているのだろうか、ということだね。それはそれなりに、オッズ比の活用法があるからだよ。



ポイント5:オッズ比の活用方法

菅:不整脈になる要因は色々あると思うけれど、ここでは喫煙有無、飲酒有無、性別を取り上げ、どの要因が不整脈の有無に影響を及ぼしているかを仮に調べることにしよう。表7はそれぞれの要因についての分割表、リスク比、オッズ比を示したものだ。

菅:まず、リスク(割合)の差を求めてくれるかな。

志賀:喫煙有無は80%－30%より50%、飲酒有無は26.7%、性別は3.3%です。

菅:差が大きい要因ほど、不整脈の影響要因といえるよ。

志賀:はい。

菅:リスク比の順位はどうなっている? 値が高い順から見てみようか。

志賀:喫煙有無が 2.67 で1位、飲酒有無が 1.67 で2位、性別が 1.07 で3位です。

表7 多数ある影響要因の比較

	不整脈		横計	リスク
	ある	ない		
喫煙	12	3	15	80.0%
非喫煙	3	7	10	30.0%
オッズ	4.00	0.43		

リスク比	2.67
オッズ比	9.33

	不整脈		横計	リスク
	ある	ない		
飲酒	10	5	15	66.7%
非飲酒	4	6	10	40.0%
オッズ	2.50	0.83		

リスク比	1.67
オッズ比	3.00

	不整脈		横計	リスク
	ある	ない		
男性	8	7	15	53.3%
女性	5	5	10	50.0%
オッズ	1.60	1.40		

リスク比	1.07
オッズ比	1.14

菅:じゃあ、オッズ比の順位は?

志賀:喫煙有無は 9.33、飲酒有無は 3.00、性別は 1.14 で、順位はリスク比と同じです。

菅:そうだね。不整脈に影響を及ぼす要因の順位は、リスク(割合)の差、リスク比、オッズ比、どれを用いても把握できるということだよ。

志賀:リスク比、オッズ比どちらを使っても、不整脈有無の影響要因の1位は喫煙有無、2位は飲酒有無ということですね。

菅:そのとおり。私はリスク比が使いやすので、リスク比で解析することが多い。でも臨床研究をする人はオッズ比を用いる人たちの方が多いように思われるね。影響要因の順位が分かれば目的を達成できるからだと思うよ。オッズ比は順位の把握だけなので、オッズ比を使う場合は有意差検定(後で解説)を併用することが必須になるんだ。それじゃ、このあたりで一度、志賀さん、今までのことをまとめてみてくれるかな。復習だ。

志賀:はい。リスク比とオッズ比の順位は同じになります。影響要因の順位を把握する目的であればリスク比、オッズ比のどちらを使っても解決できます。リスクの倍率を比較したい場合は、リスク比を適用します。表7の一番上の表の喫煙有無と不整脈有無の関係性について、オッズ比で解析した場合、オッズ比の値から「喫煙者が不整脈となるリスクは非喫煙者に比べ 9.33 倍である」と言うてはいけません。

菅:一言でまとめればどうなるかな。

志賀:オッズ比の活用は、複数の影響要因の寄与順位を把握することのみに活用できる。リスクの倍率把握には適用できない。

菅: ベリーグッドだ。

志賀: おおー！ オッズ比もゲットしました。(思わずガッツポーズ！)

菅: 少しだけ補足しておこうか。表8の分割表のリスク比、オッズ比を見てくれるかな。表7の表との違いが分かるかな？

志賀: うーん、喫煙と非喫煙のデータを入れ替えています。飲酒と非飲酒、男性と女性も同様です。

菅: 一番上の分割表について解釈してくれるかな。

表8 多数ある影響要因の比較

	不整脈		横計	リスク
	ある	ない		
喫煙	3	7	10	30.0%
非喫煙	12	3	15	80.0%
オッズ	0.25	2.33		

リスク比	0.38
オッズ比	0.11

	不整脈		横計	リスク
	ある	ない		
飲酒	4	6	10	40.0%
非飲酒	10	5	15	66.7%
オッズ	0.40	1.20		

リスク比	0.60
オッズ比	0.33

	不整脈		横計	リスク
	ある	ない		
男性	5	5	10	50.0%
女性	8	7	15	53.3%
オッズ	0.63	0.71		

リスク比	0.94
オッズ比	0.88

志賀: リスク(割合)は、喫煙者の方が非喫煙者に比べ小さくなっています。

喫煙者が不整脈になるリスクは30%で、非喫煙者のリスク80%より50%低いです。

リスク比が0.38(30%÷80%で算出)ということから、喫煙者が不整脈となるリスクは非喫煙者に比べ0.38倍です。

菅: 表7は「喫煙者は非喫煙者に比べ不整脈になりやすいという事例」だね。表8は、こんなことは実際にはあり得ないことだけど、「喫煙者は非喫煙者に比べ不整脈になりにくいという事例」となるね。

志賀: (ふむふむ)

菅: 喫煙と不整脈の関連性を見ると、表7は「喫煙あり→不整脈あり」、「喫煙なし→不整脈なし」との通常考えられる関連だけど、表8は「喫煙あり→不整脈なし」、「喫煙なし→不整脈あり」という通常あり得ない関連となっているんだ。

志賀: そうですね。

菅: 前者の関連を(表7)を正の相関、後者(表8)を逆相関ともいうよ。

志賀:先生、面白い法則を発見しました。

菅:ほおー、それは何かね。

志賀:リスク比、オッズ比どちらも、正の相関の場合は1より大きく、逆相関の場合1より小さくなっています。

菅:素晴らしい、よい発見をしたね。ここまでのところをまとめてみよう。リスク比、オッズ比ともに値が1より大きくなるほど、喫煙者は非喫煙者に比べ不整脈になる傾向が高まるといえる。このような関連性を正の相関関係があるという。リスク比、オッズ比ともに値が1より小さくなるほど、喫煙者は非喫煙者に比べ不整脈にならないという傾向が高まるといえる。このような関連性を逆相関という。

志賀:ということは、表7の場合は、リスク比は $2.67 > 1$ で正の相関、すなわち、喫煙する人ほど不整脈になりやすいといえます。表8の場合は、リスク比は $0.38 < 1$ で逆相関、すなわち、喫煙する人ほど不整脈になりにくいということです。

菅:バッチリです。

志賀:理解できましたけど、リスク比が 0.38 倍というのが何か気になるというか、わかりにくいです。

菅:因果関係としてはあり得ない表8の一番上の表の喫煙と非喫煙の位置を入れ替えた表を作り、リスク比を計算してくれるかな。

志賀:(カリカリカリ)計算しました。

表 8(2)喫煙と非喫煙を入れ変えた表

	不整脈症状		横計	リスク
	ある	ない		
非喫煙	12	3	15	80.0%
喫煙	3	7	10	30.0%
オッズ	4.00	0.43		

リスク比	2.67
オッズ比	9.33

菅:リスク比は1を上回ったね。リスク比を解釈するとどうなるかな。

志賀:非喫煙者は喫煙者に比べ 2.67 倍、不整脈になるといえます。

菅:表8の解釈、「喫煙者は非喫煙者に比べ 0.38 倍、不整脈になる」と同じことだが、上記の表現の方がわかりやすいね。リスク比が1を下回った場合、このような対応をお勧めするよ。

志賀:よくわかりました。



ポイント6:コホート研究とケースコントロール研究

菅:臨床研究にはコホート研究とケースコントロール研究があるんだけど、知っているかな。

志賀:えーと、名前はよく目にします。

菅:コホート研究とケースコントロール研究は、後者の研究で集めたデータを解析する場合、リスク比は不可、オッズ比は可だといわれている。

志賀:・・・(汗)

菅:ちょっと難しい顔しているね。じゃあ、このことを説明する前に、臨床研究の代表的研究である、「コホート研究」と「ケースコントロール研究」とは何かを説明しておこう。

志賀:ぜひお願いします。

菅:臨床研究は「前向きか後ろ向きか」で分けることができ、コホート研究は前向きの研究、ケースコントロール研究は後ろ向きの研究とされている。「前向き」「後ろ向き」の違いは、未来へ向かって調べるか、過去へ向かって調べるかの違いだね。

志賀:はい。

菅:よく理解してほしいので、今から具体例で説明しよう。

コホート研究の例

喫煙有無と不整脈有無の関連性を調べたいとする。

不整脈がない人をランダムに400人抽出し、今までに喫煙をしたことがあるかどうかを調査した。その後の2年間において、喫煙有無別に不整脈が発生したかを追跡調査した。調査開始時点では不整脈は発生しておらず、それから2年後(未来)で不整脈の発生を調べている。このような研究をコホート研究という。この研究は2年後の未来へ向かって調べる研究であり、「前向き」の研究である。

ケースコントロールの例

喫煙有無と不整脈有無の関連性を調べたいとする。

不整脈があると診断された200人とランダムに選んだ健常者200人について、過去の喫煙有無を調査した。

既に不整脈があると診断された人と健常者がいて、その時点から過去にさかのぼって喫煙をしていたかどうかを調べている。このような研究をケースコントロール研究という。この研究は過去へ向かって調べる研究であり、「後ろ向き」の研究である。

志賀:よくわかります。

菅:さっき、「コホート研究とケースコントロール研究は、後者の研究で集めたデータを解析する場合、リスク比は不可、オッズ比は可だといわれている」と話したよね。

志賀:はい。

菅:因果関係には原因と結果がある。喫煙有無が原因変数で不整脈が結果変数だ。コホート研究は未来の結果変数(不整脈有無)を調べる研究、ケースコントロール研究は過去の原因変数(喫煙有無)を調べる研究だよ。では、とあるデータを例に解説しよう。頑張って、ついてきて。



ポイント7:ケースコントロール研究ではなぜオッズ比が使われるか。

菅:次のデータはケースコントロール研究で集めたデータだよ。

表9 ケースコントロール研究で集めたデータの分割表

	不整脈		横計	リスク
	ある	ない		
喫煙	94	74	168	56%
非喫煙	106	126	232	46%
調査数	200	200	400	
オッズ	0.89	0.59		

リスク比	1.2
オッズ比	1.5

菅:喫煙者の不整脈になるリスクは56%だね。この数値から、一般的に喫煙する人の不整脈となるリスクは5割を超えているとってよいだろうか。

志賀:えーっと、ダメだと思います。

菅:なぜ?

志賀:この事例はケースコントロール研究で集めたデータです。不整脈があると診断された200人とランダムに選んだ健常者200人について、過去の喫煙有無を調査したものです。したがって、全人数による不整脈のリスク(割合)は200/400の50%で、調査対象者のサンプリング(抽出)に依存してしまいます。サンプリングに依存してしまうリスクを用いてリスク比を計算するのは間違っていると思います。

菅:そのとおり。ただ、リスク比の値を順位で検討するだけであれば、リスク比を使用してもかまない。この制限のもとにリスク比を使うぐらいなら、最初からオッズ比を使えばよいことになる。このような理由から、ケースコントロール研究の場合はオッズ比を用いるんだ。

志賀:わかった気がします。



ポイント8:ロジスティック回帰分析とは何だろう

菅:ロジスティック回帰分析とは何かを表10のデータで説明しよう。

志賀:えっ?!、ロジスティック回帰……。

菅: わからない用語だからといって難しく考えない、考えない、今からきちんと説明するからね。

志賀: よろしく願います。

菅: まず、下の表を見てみて。結果に対していくつかの原因が考えられる場合、それぞれの原因が結果にどの程度影響を及ぼしていると思う？

表 10 ロジスティック回帰説明のためのデータ

No.	不整脈 有無	喫煙有 無	飲酒有 無	ギャンブル 嗜好
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	0	1
5	1	1	0	1
6	1	1	0	0
7	1	0	1	0
8	1	0	0	0
9	0	1	0	1
10	0	1	0	0
11	0	0	1	0
12	0	0	1	0

No.	不整脈 有無	喫煙有 無	飲酒有 無	ギャンブル 嗜好
13	0	0	0	1
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	0	0	0	0

1	有り	有り	有り	好き
0	無し	無し	無し	嫌い

志賀: ひと目みただけでは、わかりません。

菅: 志賀さん、結果は「不整脈の有無」だよ。原因は喫煙、飲酒、ギャンブル嗜好の3つと想定してるね。これまでの勉強を思い出して、ちょっと面倒だけど、このデータの分割表、オッズ比、リスク比を求めてくれるかな。

志賀: はい。(カリカリカリ、カリカリカリ)

Excel の得意な志賀さん、パソコンを取り出して計算した。

表11

	不整脈症状		横計	リスク
	ある	ない		
喫煙	6	2	8	75.0%
非喫煙	2	10	12	16.7%
オッズ	3.00	0.20		

リスク比	4.50
オッズ比	15.00

	不整脈症状		横計	リスク
	ある	ない		
飲酒	4	2	6	66.7%
非飲酒	4	10	14	28.6%
オッズ	1.00	0.20		

リスク比	2.33
オッズ比	5.00

	不整脈症状		横計	リスク
	ある	ない		
ギャンブル好き	5	2	7	71.4%
ギャンブル嫌い	3	10	13	23.1%
オッズ	1.67	0.20		

リスク比	3.10
オッズ比	8.33

菅: 志賀さん、理解力もすごいけど、Excel も得意なんだ。

志賀: えへへ、おほめにいただいて光栄です。

菅: 結果を解釈するとどうなるかな。

志賀: 不整脈に影響度の最も強いのは喫煙有無、次にギャンブル嗜好です。飲酒有無は3番目でした。ギャンブル嗜好の方が飲酒有無より影響度が高いのは、何かへんな気がします。

菅: この結果が事実かどうかは、次のことを調べればわかるよ。

志賀: 何を調べればよいですか。

菅: 喫煙有無、ギャンブルの有無、飲酒の有無のそれぞれの原因の関係を見てあげるんだ。

志賀: ???

菅: 不整脈という結果に対する原因がいくつか考えられる、ということだね。これは「不整脈を説明する変数がいくつかある」ということだ。原因同士の関係性を調べてみようか。「原因要因(影響要因)相互の分割表」というものを作成し、リスク比を求めるよ。

志賀: ……はい。

菅: 飲酒有無と喫煙有無、ギャンブル嗜好と喫煙有無、ギャンブル嗜好と飲酒有無で原因要因相互の分割表を作ってみて。

志賀: (カリカリカリ)この結果でよいでしょうか…。

表 12 原因要因相互の分割表

表12

	喫煙	非喫煙	横計	リスク
飲酒	3	3	6	50.0%
非飲酒	5	9	14	35.7%
オッズ	0.60	0.33		

リスク比	1.40
オッズ比	1.80

	喫煙	非喫煙	横計	リスク
ギャンブル好き	6	1	7	85.7%
ギャンブル嫌い	2	11	13	15.4%
オッズ	3.00	0.09		

リスク比	5.57
オッズ比	33.00

	飲酒	非飲酒	横計	リスク
ギャンブル好き	3	4	7	42.9%
ギャンブル嫌い	3	10	13	23.1%
オッズ	1.00	0.40		

リスク比	1.86
オッズ比	2.50

志賀: なるほど、飲酒と喫煙、ギャンブルと喫煙、ギャンブルと飲酒のそれぞれの関係性や影響度合いを見ているんですね。

菅: そうだよ。この結果を見て、どんなことがいえるかな。

志賀: ギャンブル好きは7人中6人が喫煙者、ギャンブル嫌いは13人中11人が非喫煙者です。リスク比も5.57と高く、両社に強い関係性を感じます。飲酒有無と喫煙有無のリスク比は1.40、ギャンブル嗜好と飲酒有無のリスク比は1.86とそれほど大きな値でなく、関係性は弱いように思われます。

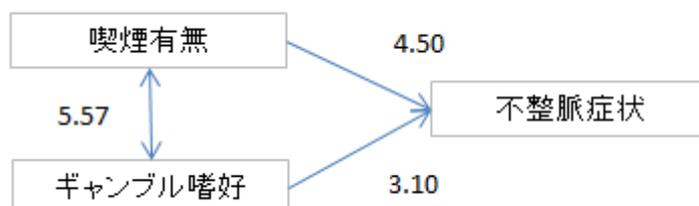
菅: そうだね。

志賀: 先生、感覚的にリスク比5.57というと、強い関係性がある気がするのですが、評価する際の基準はあるのでしょうか。

菅: 結論から言えば、統計学的な基準はない。いくつ以上の値であれば関係性が強い、言い換えれば「リスク比が●以上の値だから、強い相関がある」というのはない。リスク比同士を比較して、“強い”“弱い”と評価していいよ。

菅: よし、リスク比で、原因2つと結果(不整脈有無)の関係を見てみようか。原因は喫煙有無とギャンブル嗜好にして、図式化してみるよ。

表 13 原因要因相互のリスク比相関図



菅:喫煙するからギャンブルが好きなのか、ギャンブルが好きだから喫煙するのか因果関係の方向は分からないが、両者の関係は強いね。そして、ギャンブル嗜好と不整脈の関係が強いのは、ギャンブル嗜好が喫煙有無の影響を受けているからだと考えられる。

志賀:むむむ、なるほど。

菅:このため、喫煙有無の影響を除外したうえで、ギャンブル嗜好と不整脈との関係を調べる必要があるね。これを「真の相関関係」というよ。これを解決してくれる解析手法がロジスティック回帰なんだ。

志賀:・・・(わかるようで、わからないなあ)、ロジスティック回帰ってすごそうな解析手法ですね。

菅:計算は複雑なので残念ながら Excel ではできないよ。私が開発したソフトウェアで解いてみよう(ソフトウェアはアイスタット社から1か月間、無料貸出します。詳細はこちら)。

菅先生が解析した結果

表 14 ロジスティック回帰の結果

変数名	回帰係数	オッズ比	オッズ比95%CI		有意差判定	
			下限	上限	p値	判定
喫煙有無	2.643	14.053	0.659	299.678	0.090	[]
飲酒有無	1.948	7.018	0.457	107.843	0.162	[]
ギャンブル嗜好	0.504	1.655	0.079	34.826	0.746	[]
定数	-2.374				0.028	

菅:ロジスティック回帰を行うと、オッズ比が出力されるんだ。オッズ比から順位が把握できるので、不整脈の原因要因の1位は喫煙有無で、次に飲酒有無となる。ギャンブル嗜好は不整脈にそれほど影響がないことが分かるね。

志賀:なるほど。表 14 の結果と表 11 の結果が異なるということは、原因要因が多数あるときは、ロジスティック回帰を使わなければいけないということですか。

菅:原因要因が相互に無関係と解釈できれば、分割表のオッズ比とロジスティック回帰のオッズ比の順位は同じになる。だから、まず表 12 で示した原因要因相互の分析をして強い相関がないことがわかれば、ロジスティック回帰分析をする必要はないと思うね。

志賀:わかりました。先生、表 14 にp値がありますが、前に教えてもらった p 値と理解して良いのでしょうか。

菅:そうだよ。p値が 0.05 以下であれば、今回のサンプル 20 人から、何十万人という母集団についても、その原因は不整脈に影響を及ぼすと判断される。表 14 は、3要因ともp値は 0.05 を上回っているのだから、母集団においてこれら3要因は原因要因であるか分からないということだ。

志賀:オッズ比が 14.053 とかなり大きいのにp値は 0.05 を下回らないのは不思議です。

菅:良いところに気づいたね。それはサンプルが 25 人と少ないからだね。統計学の解析手法は、

少ないサンプルからは有意差判定ができないと判断したんだ。

志賀: 納得です。リスク比、オッズ比はゲットしましたが、ロジスティック回帰については少し不安です……。

菅: ロジスティック回帰は、原因要因相互の関係を考慮してオッズ比を算出する解析手法、くらいと理解してくればいいよ。

志賀: はい、そうします。ところで分割表から求められたリスク比から、母集団についての影響がある、ないということは言えるのでしょうか？

菅: なかなか、いい質問だね。じゃあ、その説明をしよう。



ポイント9: リスク比からの有意差検定とは

菅: 次の分割表を例に、喫煙と非喫煙のリスク(割合)に差があるかを有意差検定で調べてみようか。

表15 カイ2乗検定による有意差検定のための分割表
実測度数

	不整脈有無		横計
	ある	ない	
喫煙	60	20	80
非喫煙	50	70	120
縦計	110	90	200

志賀: はい、よろしくお願いします。

菅: では、志賀さん、まずはおさらいだ。リスクやリスク比を算出してみられるかな。

志賀: (カリカリカリ) できました！ 喫煙者の不整脈のリスクは 75%、非喫煙者の不整脈リスクは 42%で、リスク比は $75 \div 42$ で 1.8 になりました。

表15

実測度数

	不整脈有無		横計	リスク
	ある	ない		
喫煙	60	20	80	75%
非喫煙	50	70	120	42%
縦計	110	90	200	

リスク比 1.8

菅: 正解。志賀さん、この表の各セルについて、次の計算をしてくれるかな。「縦計 × 横計 ÷ 全数」だ。

志賀:「ある」「喫煙」のセルは、「ある」の縦計は 110、「喫煙」の横計は 80、「全数」は 200 なので、
 $110 \times 80 \div 200 = 44$ 。

菅: そうだね。その調子でほかも求めてみて。

志賀: はい。(カリカリカリ)できました。

表16

	期待度数 = 縦計 × 横計 ÷ 全数			
	不整脈有無			
	ある	ない		
喫煙	$110 \times 80 \div 200$	$90 \times 80 \div 200$		
非喫煙	$110 \times 120 \div 200$	$90 \times 120 \div 200$		
喫煙	44	36	80	55%
非喫煙	66	54	120	55%
縦計	110	90	200	

リスク比 1.0

菅: 正解。この計算で求められた値を「期待度数」というよ。これに対し、表 15 の人数を「実測度数」というんだ。でも、この言葉は覚えなくていいからね。

志賀: そうなんですか。しかし先生、この計算で何をしたのか、何がわかるのかは気になります。

菅: そうだよ。実測度数は調査結果のことだ。これはいいよね。で、先ほど計算してもらった期待度数だが、表 15 の喫煙者で不整脈ありは 60 人との結果だけど、60 人が多いのか少ないのか、有意な差のある数値なのかどうかはわからない。何か基準があるよね。表 16 は、この基準となる表を作ったわけだ。

志賀: ほー。

菅: 志賀さん、表 15(実測度数)と表 16(期待度数)で何か気づいたことはない？ 縦計、横計とか。

志賀: あっ！ 2つの表の値が同じです。

菅: そうだね。期待度数から求められたリスクが重要なんだけど、志賀さん、計算してみて。

志賀: (カリカリカリ)喫煙、非喫煙のリスク(割合)は 55%で同じになっています。

菅: そうなんだ。実は、実測度数表の縦計、横計はそのままの値で、喫煙と非喫煙のリスクが同じになるような4つのセルの値を求めたんだ。先ほど、「基準がある」と説明したよね。要は、不整脈の割合が喫煙と非喫煙で差がない分割表を作ったのだよ。

志賀: なんとなく理解できたような気がします。

菅: 大事なのは、実測度数が期待度数に近い値になっていれば有意差がない、離れていれば有意差があるという考えなんだ。

志賀: たしかに、そんな気がします。

菅: そこで、次の計算をするんだ。実測度数と期待度数の差を求める。求められた値を2乗して期待度数で割る。

志賀: よく分からないけれど、計算してみます。

表17 (実測度数－期待度数)の2乗÷期待度数

	不整脈有無	
	ある	ない
喫煙	$(60-44)^2 \div 44$	$(20-36)^2 \div 36$
非喫煙	$(50-66)^2 \div 66$	$(70-54)^2 \div 54$
喫煙	5.82	7.11
非喫煙	3.88	4.74
合計	$5.82+3.88+7.11+4.74=21.5$	

菅: 求められた値を合計してください。

志賀: $5.82+3.88+7.11+4.74=21.5$ です。

菅: この合計した値をカイ2乗値というよ。

志賀: この値は何なんですか。

菅: そうだね、まずはこの値がどのような時小さくあるいは大きくなるかを調べてみよう。

志賀: お願いします。

菅: 調査結果が表Aのようになったとしよう。この表のカイ2乗値を求めてください。

表A

	実測度数		横計	リスク
	不整脈有無			
	ある	ない		
喫煙	50	30	80	63%
非喫煙	75	45	120	63%
縦計	125	75	200	

志賀: 求めました。カイ2乗値は0となりました。

表B

期待度数 = 縦計 × 横計 ÷ 全数

	不整脈有無		
	ある	ない	
喫煙	$125 \times 80 \div 200$	$75 \times 80 \div 200$	
非喫煙	$100 \times 120 \div 200$	$45 \times 120 \div 200$	
喫煙	50	30	80
非喫煙	75	45	120
縦計	125	75	200

(実測度数 - 期待度数)の2乗 ÷ 期待度数

	不整脈有無		
	ある	ない	
喫煙	$(50 - 50)^2 \div 50$	$(30 - 30)^2 \div 30$	
非喫煙	$(75 - 75)^2 \div 75$	$(45 - 45)^2 \div 45$	
喫煙	0.00	0.00	
非喫煙	0.00	0.00	
合計	0.00 + 0.00 + 0.00 + 0.00 = 0		

菅: そうだね。喫煙と非喫煙のリスクがどちらも63%で同じだよ。リスクが同じ場合、カイ2乗値は0となるんだ。両者のリスクに違いがあるほどカイ2乗値は大きくなるんだ。

菅: カイ2乗値計算方法は忘れて構わないが、両者のリスクに違いがあるほどカイ2乗値は大きくなることは忘れないでください。

志賀: はい。

菅: それともう一つ、カイ2乗値は統計学が決めた3.84より大きければ、母集団におけるリスク(割合)に差があると判断することも覚えておいてください。

志賀: はい、わかりました。

菅: カイ2乗値はp値に変換することができる。p値はカイ2乗値が大きくなるほど小さくなる関係があるんだ。カイ2乗値が3.84の時、p値は0.05で、カイ2乗値が3.84より大きくなるにつれ、p値は0に近づく。

志賀: だから、p値が0.05より小さければ、母集団のリスク(割合)に差があるといえるのですね。

菅: 表15におけるカイ2乗値は21.5、この値に対するp値は0.0000だよ。したがって、母集団における不整脈のリスク(割合)は喫煙と非喫煙で差があるといえるんだ。

菅: ちなみにひとつ問題。表18は表15の人数を10で割った表だよ。

表18

サンプルサイズが小さい実測度数

	不整脈有無		横計	リスク	リスク比
	ある	ない			
喫煙	6	2	8	75%	1.8
非喫煙	5	7	12	42%	
縦計	11	9	20		

志賀: 人数は少なくなりましたが、リスクの値、リスク比は表 15 と同じですね。

菅: この表について有意差検定したらどうなると思う？

志賀: 表 18 と表 15 のリスク比は同じだから、有意差判定の結果も同じだと思います。

菅: ブー。それは間違いだよ、結果を見てごらん。カイ2乗値は 2.15 で、表 15 の 21.5 より小さくなっているだろう。カイ2乗値はリスクの差だけでなく、サンプルサイズの影響も受けるんだ。サンプルサイズが小さいと有意差がでにくいということ。ちなみに、表 18 の p 値は 0.1412 で 0.05 より大きいので、サンプルサイズ 20 人から母集団についてのリスクの差はわからないということだね。

志賀: なんとかカイ2乗検定もゲットしました。先生、ありがとうございました！

菅: 1つ言い忘れたことがあるんだ。分割表において、行数が 2 つ、列数が 2 つ(カテゴリー数が 2×2)の場合イエツの補正をした方が良いと言われているんだ。次の資料を参考にしてください。

●イエツの補正●

分割表の行数、列数が 2 つの場合、カイ 2 乗値は次の公式によっても求められます。

カテゴリー	B_1	B_2	計
A_1	a	b	y_1
A_2	c	d	y_2
計	x_1	x_2	n

$$\text{カイ2乗値} = \frac{n(ad - bc)^2}{x_1 \times x_2 \times y_1 \times y_2}$$

表 18 の分割表について求めてみます。

表 18 の再掲

表18

	不整脈有無		横計	割合	リスク比
	ある	ない			
喫煙	6	2	8	75%	1.8
非喫煙	5	7	12	42%	
縦計	11	9	20		

表 18 のカイ2乗値は

$$\text{カイ2乗値} = \frac{20 \times (6 \times 7 - 2 \times 5)^2}{11 \times 9 \times 8 \times 12} = \frac{20,480}{9,504} = 2.15$$

ただし、表 18 のように a 、 b 、 c 、 d のうちどれかが5未満の場合は、次の補正を行います。

$$\text{カイ2乗値} = \frac{n(|ad - bc| - n/2)^2}{x_1 \times x_2 \times y_1 \times y_2} \quad \text{但し、}|ad - bc| \text{ は絶対値。}$$

これをイェツの補正といいます。

$$\text{カイ2乗値} = \frac{n(|ad - bc| - n/2)^2}{x_1 \times x_2 \times y_1 \times y_2} = \frac{20(|42 - 10| - 20 \div 2)^2}{11 \times 9 \times 8 \times 12} = \frac{20(32 - 10)^2}{9,504} = \frac{9,680}{9,504} = 1.0185$$

カイ2乗値 1.0185 に対するp値を求めると 0.3129 となります

表 18 について、「喫煙・非喫煙」と「不整脈あり・なし」は関連しているとはいえない。すなわち喫煙が疾患への影響があったといえません。

志賀さんからドクターへの回答

今回勉強した知識を持って、後日、例の循環器専門医を訪ねた志賀さん。医師からの「リスク比とオッズ比、そして有意差について教えてね」との宿題に対して、

表 19 テーマの結果

	不整脈		横計	リスク
	ある	ない		
喫煙	48	143	191	25%
非喫煙	39	220	259	15%
オッズ	1.2	0.7		

リスク比	1.67
オッズ比	1.89

カイ2乗値	7.15
p値	0.0075

志賀：喫煙と非喫煙で不整脈となるリスクは、先生の計算の通り、喫煙が 25%、非喫煙が 15% で、その差は 10% です。リスク比は $25 \div 15$ から、1.67 となりました。これは、喫煙者は非喫煙者に比べ、不整脈のリスクが 1.67 倍である、と解釈できます。

ご質問いただいたオッズ比の前に、有意差についてご説明します。この 10% のリスクの差が母集団についてもいえるかどうかをカイ2乗検定で調べた結果、p 値は 0.0075 で、0.05 を下回ったので、不整脈は喫煙と非喫煙で有意な差があると判断できます。

次にオッズ比ですが、オッズは不整脈という“結果”に対して、喫煙などの“原因”がどの程度関与しているのかを分析するためのものです。不整脈がある人のオッズは、喫煙 48 ÷ 非喫煙 39 で 1.2、同様に不整脈がない人のオッズは 0.7。これらからオッズ比は $1.2 \div 0.7$ で 1.89 でした。

リスク比と同様に、オッズ比も、値が大きいほど不整脈と喫煙の関係が高くなるといえます。このため 1.89 は関係が高いと解釈できます。

しかし、オッズ比から、喫煙者は非喫煙に比べ 1.89 倍不整脈になりやすい、とは解釈できませんのでご注意ください。

オッズ比はあくまで、喫煙などの“原因”がどの程度、“結果”である不整脈に影響しているか、特に喫煙だけでなく飲酒を含めた複数の要因と不整脈との関係をみたい場合、それぞれにオッズ比を算出することで、不整脈に対する影響度合いの順位を把握することに使います。ちなみに順位の把握はリスク比でもできますので、順位把握であればリスク比、オッズ比どちらを使っても構いません。

不整脈の影響要因が多数あって、それぞれの要因が相互に大きく影響していると考えられる場合、ロジスティック回帰分析を用います。この解析手法を用いた場合はオッズ比と p 値が算出され、リスク比はありません。繰り返しとなりますが、くれぐれもオッズ比から、リスク倍率は計算しないでください。

と説明、医師からは「おお、志賀くんすごいねえ、統計が苦手なボクでもよくわかったよ。志賀くん頼りになるなあ。また今度も困ったら頼むね。」とのお言葉。とってうれしかった志賀さんだったが、次は先生からどんな難問がでるのか、ときどきでもあったのでした。