

も含めるべきか。標準化された領域を、評価を目的とする説明情報として用いることは有用である。しかしながら、どの領域が最も評価に適切かに関しては疑問が残る。また、健康状態を表すための完全な説明と単純性原則の両立の問題もある。健康状態の説明は、Rosser/Kind 行列、Health Utilities Index、EuroQol システム(表 1)など、既存の分類体系を基本とすることがあり、特定の研究のために臨時に定義されることもある。

・他のどのようなツールが健康状態を表すのに有用であるか。さまざまな研究によって、写真、漫画によるイラスト、マルチメディア・プレゼンテーションの使用など、健康状態の説明に役立つ手段が探索されている。しかしながら、追加費用、さまざまな健康状態を表す上での民族、性別、その他の要因に関する潜在的バイアスなど短所も含まれている。

各手段は、死亡に至らない健康状態の複数の領域をカバーしているが、領域や尺度に対して同じ名称を使用、あるいは各領域を構成する質問の観点から見て同じ内容をカバーしているものはほとんどない。各手段に含まれる領域の範囲は、死亡に至らない健康状態を定義および評価するためのさまざまな実証的アプローチを反映している。各領域内でカバーされている内容は多岐にわたる。手段によっては同一の領域名称を用いているが含まれている質問が異なる。各領域内でカバーされている内容の幅や深さも多岐にわたる。「視力」など特定機能を中心とした項目を含むものもあれば、「理解」や「交流」など広範囲の複雑な機能や行動を評価する項目を網羅するものもある。

表 1. 13 の一般的健康状態測定手段と領域

Health Domains (multi-dimensional profile)	QWB '70	McM '76	SIP '76	QLI '81	NHP '81	FSQ '86	CP '87	Duke '90	SF- 36 '92	HUI -III '95	WHO QOL '96	EQ 6D '99	WHO DAS II '99
<i>Overall Well-Being</i>							✓						
<i>General Health</i>				✓			✓	✓	✓				
<i>Change in Health</i>							✓						
<i>Physical Health</i>		✓						✓			✓		
Activities/roles	✓			✓			✓		✓			✓	✓
Work			✓			✓							
Home			✓										
Recreation			✓										
Ambulation			✓							✓			
Eating			✓										
Energy/vitality					✓				✓				
Dexterity										✓			
Hearing										✓			
Mobility/fitness	✓		✓		✓	✓	✓	✓	✓			✓	✓
Pain/discomfort					✓		✓	✓	✓	✓		✓	✓
Self Care			✓	✓		✓						✓	✓
Sleep/Rest			✓		✓								
Speech										✓			
Vision										✓			
<i>Social Health</i>		✓						✓			✓		
Activities/roles	✓					✓	✓		✓				
Communication			✓										
Interaction			✓		✓	✓							✓
Support				✓		✓	✓						
<i>Mental Health</i>						✓		✓	✓		✓		
Activities/roles									✓				
Alertness			✓										
Anxiety/Depression								✓				✓	
Cognition										✓		✓	
Emotional status		✓	✓		✓		✓			✓			
Outlook				✓									
Self-esteem								✓					
Understand/Interact													✓
<i>Handicap/Participation</i>					✓								✓
<i>Environmental Context</i>											✓		

QWB: Quality of Well-Being Scale, McM: McMaster Health Index, SIP: Sickness Impact Profile, QLI: Quality of Life Index, NHP: Nottingham Health Profile, FSQ: Functional Status Questionnaire, CP: COOP Charts for Primary Care Practice, Duke: Duke Health Profile; SF-36: Short-Form 36 Health Survey; HUI-III: Health Utilities Index Mark III; EQ6D: EuroQol 6 Domain Quality of Life Scale (5D excludes cognition), WHOQOL: WHO Quality of Life Bref Field Trial Version, WHODAS II: WHO Disability Assessment Schedule.

### C. 価値付けのための手法

さまざまな健康状態に対する個人の選好を引き出すために、4つの一般的手法が開発されている。

#### 1) 評価スケール/視覚的アナログスケール Rating scale/Visual Analogue scale

代表的な評価スケールは、ページ上の1本の直線と、死亡と完全な健康状態などを表

す明確に定義された 2 端点からなる。回答者は、各健康状態ですごした時間の相対評価を表すために線上の 2 端点の間に印をつけるよう求められる。

行を複数の固定カテゴリーに分割する各種の代替法や、視覚的補助手段が用いられている。このアプローチは回答を依頼する人に最も説明しやすいので、人口に基づいたデータ収集の唯一の方法となり得る。ただし結果の有効性には異論もある。特に大半の人は、極端に 0 や 1 に近い位置よりもスケールの中心に近い範囲を用いる傾向が非常に強い。

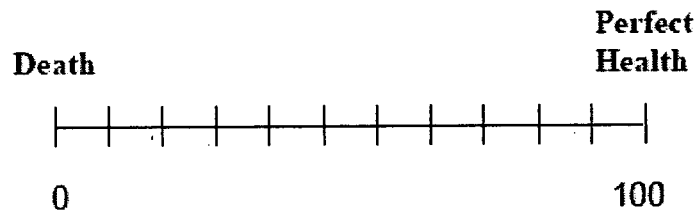


図 1. 視覚的アナログスケール

## 2) スタンダード・ギャンブル Standard gamble

スタンダード・ギャンブル(SG法)では、2つの選択肢が提供される。選択肢1では、 $t$ 年間  $1-p$  の確立で通常健康状態でいられる。死亡確率は  $p$  である。選択肢2では、 $t$ 年間慢性的に病弱な状態Aではあるが、100%の確率でそのままの状態に留まる。死亡確率がどのくらいであれば(すなわち  $p$  の値が何であれば)、2つの選択肢が気にならなくなるかというのが質問である。

健康経済学者や決定科学者によっては、スタンダード・ギャンブルは von Neumann-Morgenstern による期待効用の公理を基に公式化されているので、これが健康状態選好のゴールドスタンダードだと確信している。一方で、スタンダード・ギャンブルに対する個々人の反応には困惑が見られるという反論もある。用いられる  $p$  の値に特異的なギャンブル性(あるいは内在的リスク回避)については好き嫌いが分かれるからである。スタンダード・ギャンブルは、リスクの元での個人の健康状態選好を測定する。社会的評価は異なることがあり、個人のリスク回避を含むべきではないという主張もある。

## 3) 時間得失法 Time Trade-off

時間得失法では、状態Aでの時間  $t$ 、または状態B(通常は完全な健康状態)での時間  $x$  から選択することを求められる。2つの選択の差が気にならなくなるまで時間  $x$  が変更される。 $x$  の  $t$  に対する比率が、状態Aと比較した状態Bの相対的比重の測定値となる。 $t$  の値を 10 年とみなすことが多い。

基準期間をあまり長期(数カ月、数日ではなく数年など)にすると、個人の時間選好によって回答者に混乱が生じることになる。時間選好率が正となることによる混乱の影響は、軽い症状の場合に最も大きくなる。また、回答者の頭の中で重症度と罹患期間が明確に区別されていない時も、結果に偏向が生じる。

時間得失法は、短期間の症状の場合にも大きな問題となる。時間得失法の質問を標準化するには、比較のための良好な健康状態に対する時間  $x$  がすべての健康状態に対して同一であることが望ましい。通常は 10 年が時間  $x$  に用いられる。すなわち、回答者は 10 年間の風邪、あるいは前腕の骨折による 10 年間のギプス状態を想像する必要がある。

#### 4) 人得失法 Person Trade-off

人得失法では、健康状態 A の  $X$  人、あるいは健康状態 B の  $Y$  人のどちらを救うかの決断をしなければならないと想像する。 $X$  と  $Y$  の値がいくらであれば、選択肢の違いが気にならなくなるか。

自分自身の健康状態に対する選好は、(質問を通して明らかにされる)さまざまな健康状態にある人口集団に対する個人の社会的選好とは異なる。人によっては、自分自身のことを決断するよりも他人に関して判断する時に、生命の長さを犠牲にして生活の質を選ぶ傾向が減少する。

Nord(1995年)は、プログラム評価のためには人得失法を選択するべきだと主張する。人得失法の方が実際のリソース割り当て選択を反映しており、そのため内容に妥当性があるからである。

これに対して集団健康についての施策概要を策定するためには、区分の問題に関する個人の価値と、特定の人口集団がどの程度健康であるかという問題とを区別することが重要である。実世界のリソース割り当ては種々の選択から生じる健康のレベルのみによって決まるわけではないので、人得失法はさまざまな健康状態に対する選好の強さを測定するには最適な方法ではないという意見もある。

#### D. 回答者のタイプ

どのような人の選好を測定すべきかに関しては多くの論議がある。この問題に対する答えは、臨床判断に関する個人の価値観、あるいは集団健康の施策概要策定や社会政策立案の情報提供を目的とした社会的評価のどちらを測定するかによって大きく異なる。

#### 1) 任意の健康状態にある個人

理論上は、ある健康状態で暮らしている個々人はその状態について最も詳しい知識を持っているので、その健康状態で生きる効用 (utility) を最も正確に評価できるはずである。ところが彼らは十分な健康状態についての知識があるにも関わらず、このような場合に他の一般人とは異なる評価をすることがある。

#### 2) 医療提供者 health care providers

医療の提供者もまた、健康状態から生じる結果に関して詳しい知識を備えている。したがって、彼らも理論的には一般市民よりも確かな情報に基づいた判断を下せるはずである。ところが複数の研究によれば、医療提供者による健康アウトカムの評価は、患者の評価よりも低かった。これはおそらく、医師が日々の医療行為の中でより重篤な症例を見ていることの結果だと思われる。

#### 3) 健康な一般市民

公的リソース割り当ての判断に関しては、患者や医療提供者の価値観ではなく一般市民の価値観を用いるべきだという主張がある。一般大衆が経験したことも接したこともない健康状態について、十分な情報に基づいた選択を行えるという考えには反論の余地がある。

#### 4) 患者の家族

患者の家族は、実際に特定の病状に罹ることなく健康状態について詳細な知識を持っている。調査の結果では、家族による健康状態の評価は患者よりも低くなっている。

ある健康状態の個人から健康状態の評価を引き出す論拠の 1 つとして、適応、対処、調整の現象が挙げられている。個人による自分自身の健康状態に対する評価は、時間につれて変わることが観察されている。例えば、ある人が脊髄に傷害を負ったとする。事故のあと初めて個人が評価すると、脊髄損傷を負った後の生活の効用は非常に低くなる。しかし健康状態の選好を繰り返し測定すると、時間とともに健康状態の評価が向上することがある。

「適応 (adaptation)」とは、選択された領域において、ある人が学習や経験を通して自分の機能的性能を改善していくプロセスである。このような改善は観察することが可能であり、自覚したり報告できる可能性も高い。適応の例には、脊髄に横断損傷を負った後に車椅子で動き回ることを覚える、失明後に点字を習得するなどが挙げられる。

・人は選択領域における自覚可能な能力が観察された能力実績よりも向上するように、能力に対する自分の期待や基準を再調整することがある。このように基準を変更して、その結果自覚できる能力を向上させるプロセスは「対処(coping)」と呼ぶことができる。「対処」は人口集団内に不均一に分布していると予測されるので、同じ状況においてもある者は他のある者よりも対処に優れている。例えば、片足を失った競技者が義足を使って歩けるようになったあとで、運動性に問題がないと報告するケースがある。

・人は、自分が維持あるいは強化している機能に対するさまざまな健康状態 domain の相対評価を変えることがある。例えば、熱心なスキーヤーなど身体的活動や運動性の domain を非常に高く評価している人が、ひどい事故の後、認知機能など他の domain の価値が以前よりも上昇し、身体的 domain が低く評価されるように、domain についての自分の評価を移行することがある。domain の相対評価におけるこのような移行は、「調整(adjustment)」あるいは domain 評価シフト(domain valuation shift)と呼ばれる。

個々人の自分の健康状態に関する価値付けは、必ず適応、対処、調整を反映している。

個々人が社会的評価を行う場合には、対処や調整を取り入れない方がよい。このような valuation が用いられて健康リソースの割り当てに影響を与えると、個々人が対処、調整するという理由だけで、ある健康状態を予防したり改善するプログラムを排除する結果になる恐れがある。

#### E. 測定アプローチ

各状態を説明、評価するためのさまざまな方法に加えて、幅広い健康状態に対する評価を引き出すための全体的な測定計画の実施に関する重要な問題がある。

さまざまなレベルの重症度にわたる多様な状態を同時に評価すれば、結果の信頼性と有効性が向上することが研究によって示されている。

測定計画では、評価を必要とする状態の合計数も考慮に入れる必要がある。代表的な臨床判断分析に関しては、通常数種類の状態のみが必要とされるので、各状態に対する評価は直接引き出すことができる。これに対して集団健康に関する施策概要では、該当するすべての健康状態に対する直接評価を引き出すことは不可能である。そのため、直接評価を引き出せない状態に対して値を割り当てるための何らかのアルゴリズムを開発することが必要である。

健康状態が、複数の健康領域を基本とした標準的な記述システムを用いて表されてい

れば、さまざまな領域レベルに基づいて評価予測できるモデルを推測することが可能である。このためには、さまざまな統計や領域レベルの組み合わせから構成される一式の健康状態を表す説明を定義して、これらの状態に対する評価を引き出し、計量経済学的方法を適用して領域レベルと評価の関係を推測する。この1例が、オーストラリア疾病負担研究(Australian Burden of Disease Study)用に開発された回帰モデルであり、EQ5D+の6つの領域上のスコアを、オランダ価値付け研究(Dutch valuation study)の評価パネルから導き出された障害の重篤度と連結している。回帰方程式は以下のようになる。

$$\ln(1-DW) = \text{constant} - 0.247*d_{12} - 0.5419*d_{13} - 0.12455*d_{22} - 0.5718*d_{23} - 0*d_{31} - 0.205*d_{33} - 0.0576*d_{42} - 0.5039*d_{43} - 0.2284*d_{52} - 0.742*d_{53} - 0.1915*d_{62} - 0.5635*d_{63} - 0.1228*\text{prognosis} - 0.01*\text{disease state present} - 0.20369*\text{annualised health state}$$

定数(constant)

予後(prognosis)

疾病状態が存在(disease state present)

年換算健康状態(annualized health state)

定数 = 0; 領域1上に2のスコアが存在するか0のスコアが存在すれば  $d_{12} = 1$ ; 領域1上に3のスコアが存在すれば  $d_{13} = 1$  など。播種性腫瘍、AIDS、ひどい認知症など健康状態の予後が悪ければ予後のスコアは1となる。“disease state present”は全領域が1の値を取りなおかつ疾病状態が存在する場合に1の値を取るダミー変数である。

“annualized health state”というダミー変数が含まれた理由は、オランダの価値付け研究において短期の感染状態が変わった方法で扱われたからである(その内の2週間が肺炎である1年の健康状態を評価)。

例えば「深刻なうつ病」という健康状態はEQ5D+では223232と表される。

この健康状態に対する障害重み付け係数(disability weight, DW)は次のようになる。

$$DW = 1 - \exp(-0.247 - 0.12455 - 0.205 - 0.0576 - 0.742 - 0.1915) = 1 - \exp(-1.69055) = 1 - 0.184 = 0.816$$

評価の実行によって得られた深刻なうつ病に対する実際のDW値は、0.76であったことに留意すること。回帰モデルを用いたドイツの研究で評価されていない健康状態に対して導き出したDW値は、同じ研究会がこの健康状態を評価するように求められた場合に算出されたであろうDWの近似値と考えるべきである。もちろんこれは正確なモデルではないが、0.89という $r^2$ 値は妥当性のある近似を示している。

## VI. その他の課題

### A. フレーミング効果とアンカリング効果 Framing and Anchoring Effects

ある健康状態に対する選好は、健康状態を完全な健康あるいは死亡と比較しているかによって大いに左右される。検討の範囲は、このようなタイプの選好に関する質問への回答に対しても重大な影響を与える。種々の評価質問はそれぞれ、さまざまなフレーミング効果の影響を受ける。

### B. 信頼性と有効性 Reliability and Validity

いくつかの手法では、評価者間(inter-rater)信頼性や再検査(test-retest)信頼性が測定されているが決定的な研究はない。価値明確化の仮説に従って、各個人は従来の再検査信頼性研究で用いられている時間間隔についての反省を活かして、自分たちの評価を変更できる。

明確なゴールドスタンダードが存在しないため、健康状態選好に対する有効性は定義が困難である。

Nord は、反省的均衡を用いて有効性を測定することを提案した。選好が意味する内容を、次のような記述形式で回答者へフィードバックすることができる。

対象者は、状態Aでの2年間の生活と状態Bでの1年間の生活の差異が気にならない。対象者は、状態Aで1人の患者に2年間生活させるのと、状態Bで4人の患者に1年間生活させる場合の差異が気になるか、など。

### C. 死亡より悪い健康状態

数多くの研究に、「激しい痛みを伴う寝たきり状態」など、何人かの回答者が死亡より低い効用値 (utility) 評価をした健康状態が含まれている。

このような所見の政策的含意では、安楽死が望ましいとされることがある。

評価研究が死亡より悪い状態を認めた場合、基数値の解釈は明確ではない。死亡と完全な健康の間の状態とは異なり、死亡より悪い状態にはアンカー一点が1つしかないので、スケーリング特性を定めるのが困難となる。

実際は、これほど重篤な健康状態はまれであり期間も限られるので、疾病負担予測に



大きな割合を占めることはない。

#### D. Distribution Consideration

オレゴン州健康サービス委員会(Oregon Health Services Commission: 合衆国オレゴン州)は、地域の共同体は個人への小さな損失よりも個人への大きな損失に対してより関心があることを示した。このいわゆる「救済ルール(rule of rescue)」は、健康サービス提供の公平性に関する文献の大半に記される基本前提に逆行する。

健康年数相応(healthy year equivalent)など、重症度と罹患期間の独立という前提を必要としない代替評価アプローチが提案されている。残念ながら代替アプローチでは、重症度と期間の可能な組み合わせすべてを評価するように個人々々へ依頼する必要がある。

#### VII. GBD および NBD 研究における健康状態の価値付け

最初に発表された GBD の結果は、人得失法を採用して、国際的な公衆衛生専門家で構成される、限られた数の評価委員会の意見に基づいた健康状態価値付けを用いた。

その結果、ほぼ同様のメソッドがオランダの NBD 研究用に複製された。この研究は、EQ5D の 6 次元改良型によって評価される各健康状態を系統的に表した。6 番目の認知領域が、運動性、セルフケア、通常活動、苦痛/不快、うつ/不安という通常領域に追加された。6 つの領域にスコア 1(問題なし)、2(問題あり)、または 3(深刻な問題)を与えて、各健康状態を表した。この研究で評価されている健康状態は、死亡率が低い国々に共通する条件を反映している。大半の疾患で複数レベルの重症度が明示的に評価されるようになったことも有用な点である(軽度、中度、重度の喘息またはうつ病など)。

現在の GBD 予測は多くの NBD 研究と同様に、オリジナルの GBD とオランダの障害重篤度を混合して用いている。これは完全に満足できるアプローチとは言えない。世界疾病負担研究(Global Burden of Disease Study)2000 に向けて、WHO は次のような 2 層構造のデータ収集計画を策定した。

1. 世界健康調査(World Health Survey)の一貫として、多重状態ランキングと視覚的アナログスケール質問表を用いた一般集団における健康状態評価の収集
2. 多重状態、多重メソッドによるグループ訓練プロトコルを用いた教養のある回答者間の健康状態評価の収集

計画の最初のパートは、さまざまなタイプの回答者と環境にわたる健康状態評価に生じ

るばらつきの問題に対処することを目的としている。国内および海外においてばらつきの潜在的決定因子を求めるために、結果が分析される。

計画の第2のパートは、ランク表や視覚的アナログスケール質問表を用いて取得した回答を、選好強度を測るための区間スケール測定法へ変換するのに必要とされる。選好強度は集団健康の施策概要策定に用いることができる。

地域の健康状態評価の算定を希望する国々にも、WHOによって開発された標準化プロトコルに準拠する同様の2層データ収集計画の使用を勧めた。

しかしながら残念なことに、この研究の結果は出版されていない。この研究が直面している難題の1つは、どのような集団内でも同じ健康状態に対する評価に大きなばらつきがある点と（またこの集団内のばらつきは、集団間のばらつきよりもはるかに大きい）、さまざまな評価技術を用いた多くのレベルの回答が存在する点である。疾病負担の分野は、明らかにさらなる発展を必要としている。

---

# **DISMOD:**

## **A Tool for Internally Consistent Epidemiological Estimates**

---



### I. はじめに

Burden of Disease (疾病負担) 研究では、疾患の発症率、有病率、罹患期間、重症度、致死率の見地から疾患疫学を定量的に表すことが必要である。データの利用はこのための基本となるが、実際には次の2つの理由から疫学的データは不十分なことが多い。

- ・データが不足しているので、補足的な情報を取得して疾患を説明する必要がある。
- ・データを利用できる場合でも内部一致性に欠ける。

どちらの場合にも DisMod という特別なツールを利用すれば、専門的知識によって利用可能なデータを補足し、予測値に内部一致性を持たせることができる。DisMod には旧版の DisMod I と、機能を強化した最新版の DisMod II という2バージョンが存在する。このツールについて説明するためには、まず疾患頻度を表すために疫学で用いられる多くの概念を定義する必要がある。

### II. 比率、リスク、確率など

#### A. 粗雑な疫学定義

各種のデータソースを扱うには、比率、リスク、確率に関するある程度の基本定義を明確にすることが重要である。

1つに決められた標準定義セットは存在せず、辞書ごとに比率、リスク、確率をさまざまな方法で定義している。

Fleiss は「確率、相対頻度、割合、比率という用語は同義的に使われている」と述べている (p. 1)。このように用語の使い方が粗雑なため、初めにきちんと定義しておく必要がある。

#### B. 3つの主要な疾病頻度測定方法

疾病頻度を測定するには、主に頻度、比率、確率という3種類の方法がある。少数の例外を除けば、すべての疾病頻度測定は主要タイプのいずれかに該当する。そのため初めに、疫学分野で使われている用語を用いて3方法を定義する。

比率とは、ある時点における症例数を同時点での総人口数(症例数を含む)で割ったものであり、次の式で表す。

**Proportion = Cases/Population**

(比率 = 症例/人口)

症例数と人口はどちらも同じ一時点で測定され同じ単位(人数)で表されるので、結果として生じる比率は無次元数となる。

確率は、ある期間内の新たな症例数を、その期間の開始時においてリスクがある人数で割った値、すなわち次の式で表される。

**Probability = Number of new cases with disease during a certain period/Total number of persons at risk at the start of the period**

(確率 = ある期間中の新たな疾患症例数/期間開始時における危険人口)

分母は「リスクがある人」であり、その期間の開始時にすでに発症例に該当するなど、理由を問わずリスク状態にない人を除外している。また確率には常にそれと関連した期間が含まれるにも関わらず、無次元数となることにも留意する。

比率は、ある期間中の新たな症例数をリスクがある人時で割ったものである。

**Rate = Number of new cases with disease during a certain period/Total number of person-time at risk**

(比率 = ある期間中の新たな疾患症例数/リスクがある人-時の合計数)

分子は確率と同じになるが分母は異なる。リスクがある人-時の概念は図 4.3.1 に示されている。ここでは、ある人口集団について5年間調査した研究と4名の研究対象を例にあげている。最下部の Person 1(人物1)は研究期間の5年を通じて健康(すなわち疾患にかかるリスクがある)状態であったため、リスク状態の5人年としてデータに寄与している。Person 2は2.1年後に病気になり1年後に死亡した。病気になるとその人はもはや疾患にかかるリスク状態にはないので、2.1人年のリスク状態を構成してい

る。同様に研究期間の最後まで生存していた Person 3 は、3.7 人年のみのリスク状態を構成している。最後に Person 4 はリスク状態の 2.2 人年としてデータに寄与した後、フォローアップ不能か死亡となっている。したがって、5 年間のこの研究における 4 名に関するリスク状態の合計人時は、理論最大値 20 のうちの 13 となる。

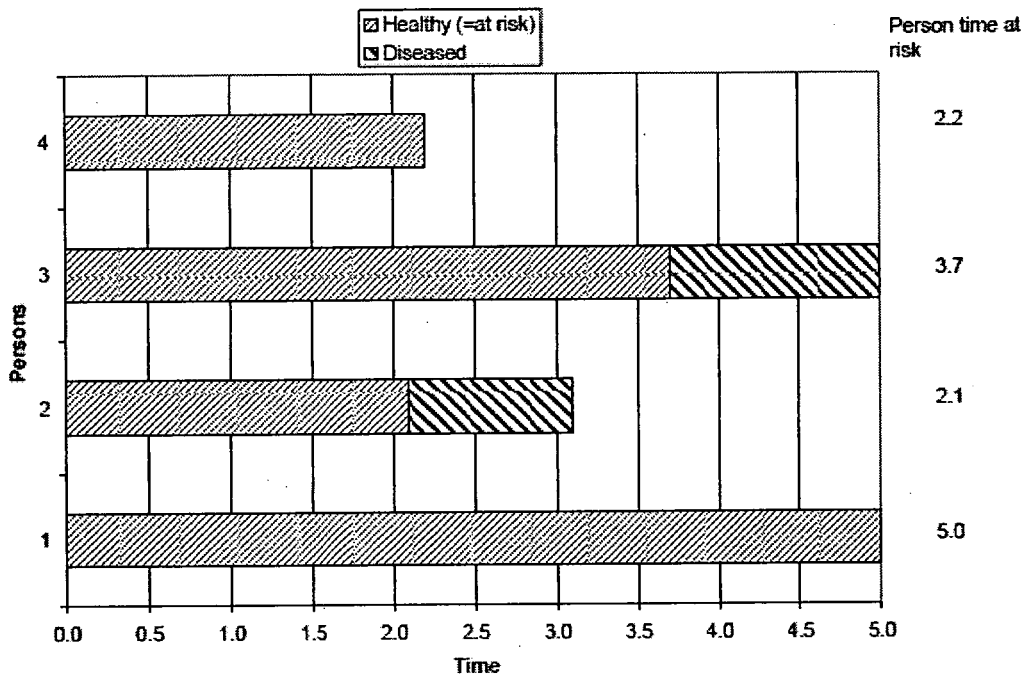


図 4.3.1: リスクがある人一時 person-time at risk

### C. 確率と比率

確率と比率は、「健康状態」から「罹患状態」、「生存」から「死亡」などある状態から別の状態への移行を表す 2 つの方法である。果たす目的と定義は同様であるが、2 つの間には重要な違いがいくつかある。

- ・ 確率とは対照的に、比率は無次元数ではなく 1 次元/時である。
- ・ 確率は 0 から 1 の値を取り、比率は 0 から  $\infty$  (無限大) である。これは、確率と比率の数値はリスクが低い時には接近しているが、リスクが上昇するにつれて差が大きくなることを意味している。

上述した人時に関する考察から、比率予測の方が確率予測よりも深く関与することが示される。すなわち、比率予測では研究期間を通じて研究対象者すべてを詳細に観察する必要があるのに対して、確率予測では研究開始時と終了時のみに観察すれば済む。

それにも関わらず比率は、疾患のモデル化においてある状態から別の状態への移行を表すために好んで用いられる方法である。この理由を探るために、初めに「競合リスク」の概念について考察する。

#### D. 競合リスク Competing risk

塗抹陽性肺結核の患者 1,000 人のコホートを 2 年間余りフォローアップしたと仮定する。コホート内の個人には、「死亡」、「回復」、「塗抹陽性肺結核に罹患したまま」の 3 つが起こり得る。

以下の図は、基本事例と仮想的な新治療法の導入後例を比較した場合の、2 年間にわたるコホートの進展を示している。どちらの例においても、時間経過とともに治癒するか死亡するコホートの割合は増加している。

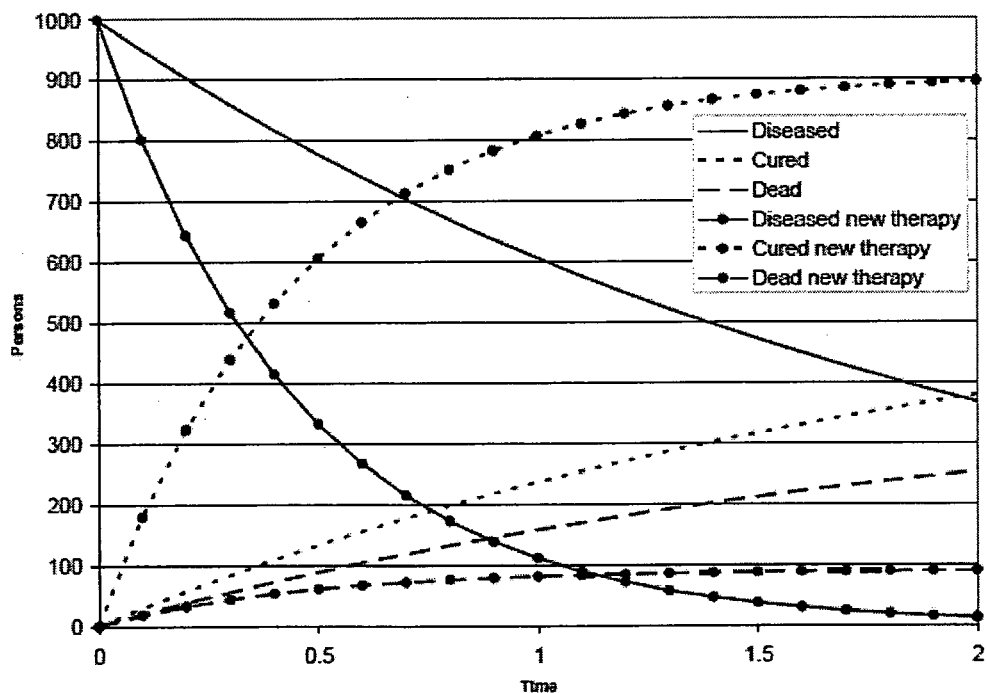


図 4.3.2: 塗抹陽性肺結核患者のコホート

考察を容易にするために 3 つの変数を定義する。 $D(t)$  は  $t$  の時点におけるコホート内の死亡数、 $C(t)$  は  $t$  の時点で治癒している研究対象者数、 $P(t)$  は  $t$  の時点でも塗抹陽性肺結核にかかっている患者数とする。

当然のことながら  $P(0)$  は観察期間開始時におけるコホート内の人数となり、筆者らの例では 1,000 人である。

コホート P 内の人はいかなる時点においても同時に 2 種類のリスクを負っている。すなわち死亡するリスクと治癒するリスクである。t の時点でまだコホート P 内にいる人数は、この 2 種類のリスクの組み合わせによって決定される。t の時点で亡くなっている人数は死亡のリスクにより決定されるが、治癒するリスクによっても決定されるのである。

このことを理解するために、コホート P 内の人々が任意の死亡リスクを負っていると仮定する。新しい治療法によって治癒のリスクが上昇すると、人々の C への移動が速くなり、しばらくするとまだ死亡リスクのある人が P 内にあまり残っていない状態になる。一方、基本事例では治癒リスクが低いためにより多くの人が P に残って死亡リスクを負い、その結果最後に D に入ることになる。複数のリスクが同時に結果の決定要因になっている時に、これらのリスクは「競合」していると言われる。

t の時点での死亡数を表す方法の 1 つは、次の式で示される死亡確率である。

$$\text{Prob}(\text{death})_t = D(t)/P(0)$$

同様に治癒確率は以下の式で示される。

$$\text{Prob}(\text{cure})_t = C(t)/P(0)$$

死亡と治癒のプロセスを測定するこれらの方法は、数学的に条件付き確率と呼ばれる。2 年後に治癒する確率は、2 年後に死亡する確率に必ず依存するので条件付きと呼ばれるのである。複数リスクを持っているという本来は継続的なプロセスが、区間開始時の人口と終了時の死亡数という 2 時点だけを見て表されるので、治癒と死亡確率の相互依存性(およびその結果による条件付き性質)が生じてしまう。

我々のコホート例では、新治療法によって治癒リスクのみが変化して、死亡のリスクは一定のままであった。ただし競合リスクが原因で、2 の時点で死亡確率は約 25%から約 9%に減少した(図 4.3.2 を参照)。

このような確率はコホート研究から容易に測定できるが、本質的に条件付きであるのでモデル化で使用することはためらわれている。ある確率が変化すると、競合確率がすべて変更されてしまうのである。

## E. 比率使用のための解決策

競合リスクの問題は、本来継続的なプロセスを不連続化することで生じる。この問題は、「リスクがある人時」を正確に測定してそこから比率を計算することにより回避できる。上述の結核事例では、比率として表された死亡リスクは一定であった。

リスクを継続的に表す比率は、通常「ハザード率」、「瞬間一率」、「一力」、または「一密度」と呼ばれる。

結核コホート事例では、治癒と死亡のハザード率を計算する前に、まずシンプルなモデルを作成する必要がある。

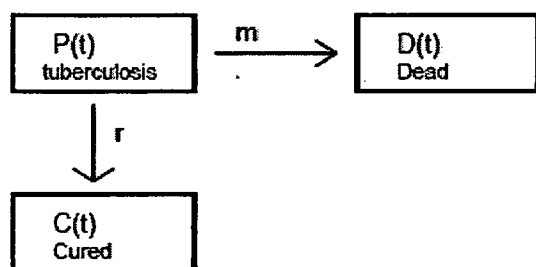


図 4.3.3: 競合リスク

3つのボックスは、「結核に罹患したまま」、「死亡」、「治癒」という起こり得る3種類の結果を表す。 $m$ は結核に罹患した個人が死亡するハザード率、 $r$ は結核に罹患した個人が治癒するハザード率である。

長期的な疫学研究からハザード率を算出することができる。正確な人時の値が測定されている場合には、新たな症例数を人時の値で割るだけで計算できる。正確な人時の値が測定されていない時には、時間区分0から $t$ の間にハザード率が一定であったと仮定できる。そうすると、結核事例に関する治癒ハザード率の式を次のように示すことができる。

$$r = \left( \frac{C(t)}{C(t) + D(t)} \right) \frac{-\ln \left( 1 - \left( \frac{C(t) + D(t)}{P(0)} \right) \right)}{t}$$



また、死亡のハザード率は次のように表せる。

$$m = \left( \frac{D(t)}{C(t) + D(t)} \right) \frac{-\ln \left( 1 - \left( \frac{C(t) + D(t)}{P(0)} \right) \right)}{t}$$

これらの方程式は事象確率の合計を合計事象率(2番目の部分)に変換し、次に合計事象率を構成要素である2つの事象間に比例的に分配する(1番目の部分)。これらの公式を用いれば、治癒および死亡の基礎プロセスのいかなる指標もコホート研究から計算できる。

#### F. 母比率

コホート研究を利用できる場合には、比率を計算することにより競合リスク問題を回避できることが実際に示された。では利用できない場合にはどうであろうか。人口動態事象記録を利用する人口統計学者や研究者は、母比率を報告することが多い、これは、ある期間に発生する事象(死亡または治癒)数を中間期人口で割ったものである。年齢別の死亡率は、1つの年齢グループに対する母比率(1年間の死亡数を年央人口で割ったもの)である。年央人口は直接観察することも、年末(または年頭)人口観察値2つを平均することによって近似値として求めることもできる。

結核患者のコホートでは、期間  $t$  にわたる類似人口死亡率を次のように計算できた。

$$m = D(t)/P(t/2) + C(t/2)$$

治癒率についても同様に計算できた。

母比率は、おそらく最も一般的な死亡率と疾病発症率の測定方法である。期間が長すぎず(通常1年)死亡ハザード率が高すぎなければ、中間期人口は妥当性のある人年数の近似値となる。

したがって、生存者全員がリスク状態である死亡率では、これらの母比率はハザード率の妥当な近似値の役割を果たす。ところが、発症していない人だけがリスク状態となる発症率の場合には、高齢での認知症などの疾病有病率が高いと母比率はハザード率から大幅に逸脱することがある。

### III. 疾病の測定と DISMOD

本セクションでは疫学で用いられる疾病頻度の測定法について考察し、通常の疫学的定義と DisMod で用いられる定義との違いに注目する。

#### A. 有病率 Prevalence

時点有病率(point prevalence)は、 $\text{Number of cases/Number of people in population}$  (症例数/人口集団の人数)と定義される。

症例と人口は同時に測定し、症例は人口に含められる。したがって有病率は割合となる(疫学者は通常「prevalence rate」と言うことが多い)。

場合によっては時点有病率の代わりに、1ヵ月有病率または1年有病率など「期間有病率(period prevalence)」が報告される。期間有病率は、

$\text{Number of people with disease during a certain period/Total number of people in the population}$  (ある期間中の罹患者数/人口集団の合計人数)と定義される。

時点有病率は、うつ病など発作性疾患を研究する疫学者の間でよく用いられるようである。時点有病率の考え方は、症例数を多くして有病率予測の信頼区間を短くすることである。

しかしながら、期間有病率には次のような多くの明確な短所がある。

- ・ 期間有病率は疾病の発症期間に依存する。
- ・ 期間中、1人が複数の疾病を発症していることがある。
- ・ 疾病の発症が原因で、死亡した症例を見逃してしまう。
- ・ 想起バイアスの可能性がある。

これは期間有病率が役に立たないことを意味するものではない。例えば、ひどいうつ病(Major Depression)に対する1ヵ月間の有病率予測は、非常に有効な時点有病率の推定量となる。

DisMod I は時点有病率を出力値としてのみサポートしており、DisMod II は入力値としても認めている。DisMod I と II のどちらも期間有病率をサポートしていない。

## B. 発症率 Incidence

発症率予測は、ハザード率 (hazard rates)、母比率 (population rates)、確率 (Probabilities) という 3 種類の形態をとる。この 3 つについて順に考察する。

発症ハザード率は次のように定義される。

Number of new cases with disease during a certain period/Total number of person time at risk

(ある期間中の新たな発症数)/リスクがある人時の合計数

発症ハザード率 (incidence hazard) は「発症密度 (incidence density)」、「発症力 (force of incidence)」とも呼ばれる。

この定義の重要な部分は「危険人口 (person time at risk)」である。人は生存中で罹病性があり、発症していない時にリスクがある状態と言える。発症率を予測するには 3 種類の方法のうち発症ハザード率が最も精度が高いが、非常に詳細な測定が必要とされるのでほとんど利用できない。

代替方法としては母比率 (population rates) がよく利用される。

Number of new cases with disease during a certain period/Average number of person in the population

(ある期間中の新たな発症数)/人口集団内の平均人数

「人口集団内の平均人数」は、リスクがある人時に対する近似値の役割を果たす。先に述べたように結果として生じる母比率は、ハザード率と疾患有病率が高くない限り発症ハザードの妥当性のある近似値となる。

発症確率 (incidence probability) は次のように定義される。

Number of new cases with disease during a certain period/Total number of persons at risk at the start of the period

(ある期間中の新たな発症数)/その期間の開始時にリスクがある人の合計数

発症確率は疫学者によって「リスク」と呼ばれることが多く、「累積発生率」、「発症比率」、「発病率」などの呼び名も使われている。

発症確率は DisMod によってサポートされていないが、発症確率を母比率に変換するこ

とができる。この変換では期間中に率が一定であると仮定する必要がある。式は以下のようになる。

$$\text{Rate} = -\ln(1-\text{Prob})/t$$

t は期間中の時間単位数である。1 年間にわたって確率が測定され年次比率を求める時には、t は 1 になるので無視することができる。

比率から確率への逆算でも定率の仮定が必要であるため、次の式を用いる。

$$\text{Prob} = 1 - \exp(-t \text{ Rate})$$

DisMod I は発症ハザード率のみをサポートしている。DisMod II は発症ハザード率と母比率の両方をサポートしているため、ユーザは用いる時に違いに注意する必要がある。母比率では全員にリスクがあると見なすのに対してハザード率では発症していない人だけにリスクがあると見なすとして計算されるので、これは特に有病率が高い場合には、任意の発症例数に対してハザード率が母比率より高いことを意味する。

### C. 死亡率

死亡率は、疾患別死亡率、合計死亡率のどちらも比率または確率として表すことができる。例えば Statistics Netherlands は合計死亡率を確率として報告しているのに対して、疾患別死亡率は人口比率として報告されている。この理由は明確ではないが、このような確率と比率の間に変換式を用いれば両者の違いによる混乱を避けることができる。

死亡率は次のように定義される。

Number of deaths during a certain period/Average number of persons in the population

(ある期間中の死亡数)/人口集団における平均人数

国家統計局による報告では死亡率は必ず母比率になる。これは総人口のレベルで人時データを得るのは不可能だからである。しかし発症率の場合と異なり、生存者のすべてにリスクがあると見なすことができるので、母比率は死亡ハザード率の極めて精度の高い近似値と言える。

DisMod は入力値として合計死亡率を必要とする。DisMod I は疾患別死亡率を出力値と