

スクリーニングテスト

スクリーニングテストとは何でしょうか？スクリーニングテストは安価で簡便な検査によって症状が出現する以前に病気を発症する可能性のある人を選び分ける手法で、早期に診断し、早期に治療を施すことによって予後を改善するのが目的です。1996年WHOが以下のように定義しています。

Scientific

Validity and reliability of test

Natural history of disease

Effective treatment available

Ethical

Does treatment prolong life?

(lead time / length bias)

Is program acceptable to target population?

Economic

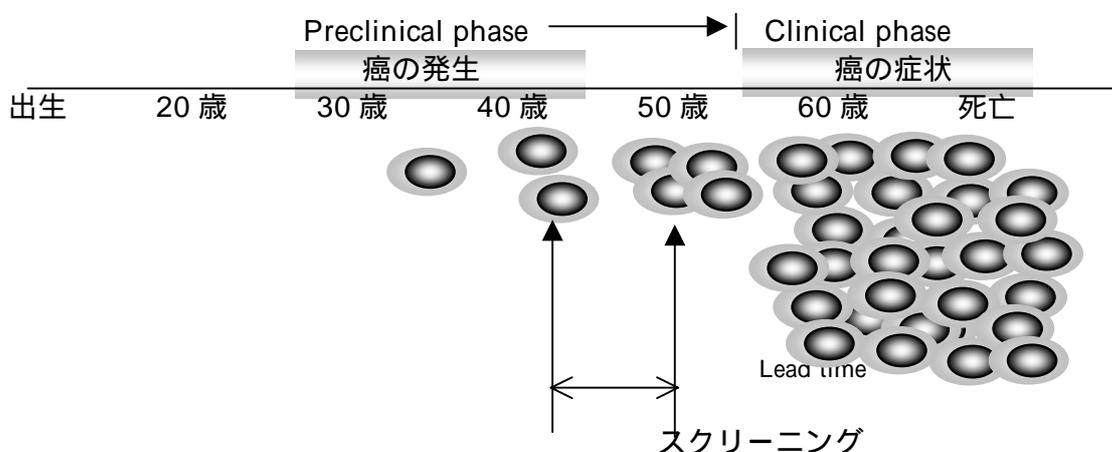
Can we afford false positivities?

Can we afford follow-up of positive patients?

Is the program cost effective?

まずは科学的に信頼性の高いテストでなくてはなりません。もしも当るも外れるも50%であればテストの意味がありません。後に述べますが感度と特異性は重要な問題です。スクリーニングテストは癌を早期の段階でとらえようとする意図があります。

Lead time bias



lead time とはスクリーニングで癌細胞が見つかったから症状が発現するまでの時間です。これは癌によってある程度わかっているものもあれば判っていないものもあります。もちろん癌の種類が同じでも急速に成長するものもあれば、ゆっくりと成長するものもあります。それでは Lead time bias とは何でしょうか？例えば上で 50 歳の時にスクリーニングで癌が発見されて 70 歳で亡くなったとすると、生存時間は 20 年です。一方その癌をもつ患者さんの natural history はスクリーニングをしていないと 60 歳で発症し 70 歳で亡くなるとします。一見 10 年しか生きられない癌をスクリーニングにより 20 年にまで延長したように見えます。これを lead time bias と呼びます。しかし、癌の成長に関しては実際何も変わっていません。

Length bias

Length bias とは何でしょうか？早期胃癌や神経芽細胞腫のスクリーニングでみられたように、非常にゆっくりと成長する癌をスクリーニングでひっかけており、本来目指していた癌をひっかけていない場合に生じます。そのゆっくりと成長する癌は死ぬまで症状を発しないか、神経芽細胞腫のように自然に消滅してしまうかもしれません。しかしスクリーニングで発見されて、病理学的に癌細胞が発見されればやはり治療してしまうのが現代医学です(だまってみるのも勇気がいりますが、勇み足ということもあります)。そして本来死ぬまで発症しない癌を治療するわけですから予後が良いのは当たり前です。そしてスクリーニングにより予後がよくなったと錯覚してしまうのが length bias です。急速に成長する癌よりはゆっくりと成長する癌の方がスクリーニングにひっかかりやすいので当然このような現象が認められてしかるべきでしょう。それでは本当にスクリーニングが有効であった場合とどう違うのでしょうか？スクリーニングで予後が改善したのであれば、その癌による死亡数(絶対数)が減少しなくてはなりません。もしも length bias を生じている場合はその癌の incidence が増え、しかし死亡数は変わらないこととなります。このような状況ではスクリーニングを中止するべきなのです。

スクリーニングテストの実際

経済的見地からしても、癌の発生ないしは死亡率を減らすことが無ければお金の無駄になります。また経済的な問題ですが、その癌がそこそこの頻度で発見されないと人々全体への効果を考えたときには価値を見出す事ができません。また最近癌と関係の深い

遺伝子変異がみつきりつつありますが、スクリーニングでこれを発見しても、有効な治療法がなければ癌を宣告しているようなもので、しかも患者さんの悩む期間が長くなる分倫理的に問題です(行なうべきではありません)。また採尿、採血程度ならまだしも、大腸ファイバーをスクリーニングで行なうとすれば人々に受け入れられないかもしれません。疑陽性が非常に多い場合、その人達に癌が無い事を確認するための更なる検査が必要となりますが、費用の点もからんできます。また癌でないのに癌かもしれないといわれた人の精神的負担を考えなくてはなりません(なかなか数値として表現できるものではありませんが、相当大きいと考えるべきです)。もちろん癌であった場合でも、早期に発見されたとは言ってもその精神的負担は大きいはずです。このような状況を考えてアメリカでは乳癌スクリーニングの前に心理相談を行なっているところもあります。また乳癌検診で false positive であった女性 2400 人について検討したところ、mammography では 18.6%で生検を施行されていますが、触診では 6.2%でした。つまり mammography は overdiagnosis になりやすい傾向にあるといえます。Mammography には通常 100 ドルの費用がかかりますが、この false positive のために更に 33 ドル支出することになります(NEJM 16;1089-96,1998)。最近では mammography をコンピュータで診断するシステムが導入されましたが、mammography の診断結果は読影する放射線医師によって随分違うようです(NEJM 131;1493-9, 1994)。

大腸ファイバーによるスクリーニング

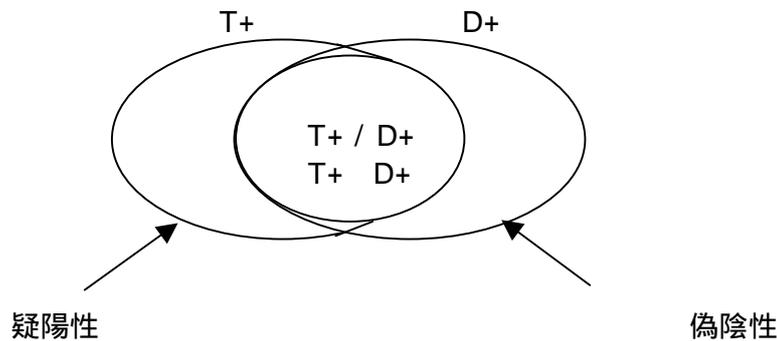
大腸ポリープは大腸癌に進展する可能性があります。National Polyp Study Workgroup は 1418 人の大腸ファイバーによりポリープ(腺腫)を切除した患者さんを平均 5.9 年 80%経過観察したところ 5 人において早期大腸直腸癌を発見したと報告しています。これは通常の大腸癌が発見される頻度より低く、ポリープ切除により癌への進展を予防できた可能性を示唆しています(NEJM 329;1977-81,1993)。Sigmoidoscopy を行なった 3496 人において 311 人(9%)にポリープを発見することができました。この患者さんのうち同意した 203 人に colonoscopy を行なったところ約 3 割に更に先に癌を見つけることができました(NEJM 336;1761-2,1997)。大腸癌において未だスクリーニングの方法に定石はありませんが、受診者の負担を考えると sigmoidoscopy、次に colonoscopy が適当かもしれません。また最近ではヘリカル CT との比較が行なわれ、癌は 3/3 (100%)、10mm 以上のポリープは 20/22、6-9mm のポリープは 33/40、5mm 以下のポリープは 29/53 でした。スクリーニングに用いるには false negative が多く十分とはいえませんが、受診者の負担を考えると魅力ある方法です(NEJM 341;1496-503,1999)。経過観察のための colonoscopy も 3 年後で十分という説もあります(NEJM 328;901-6,1993)。またリスクの高い人に対して積極的に検査するのも有効です。両親あるいは兄弟に大腸ポリープがあった場合、その人が大腸癌に罹患する relative risk は妻あるいは夫と比較して(同年代であり、似た環境で生活しているため遺伝的因子を検討する際しばしばコントロールとして用いられる) 1.78(95% CI 1.18 - 2.67)となります。家族のポリープ発見が 60 歳以前であった場合には relative risk 2.59 (95% CI 1.46 - 4.58)まで跳ね上がり、若く診断されればされる程そのリスクは高くなります(NEJM 334;82-7,1996)。大腸癌では microsatellite instability が若年発症の大腸癌に多く(17%)認められますが、これはむしろ予後良好因子として認識されています(NEJM 342;69-77,2000)。現在、大腸癌に対してさまざまなスクリーニングの試みがなされています。大腸癌死亡率を減らせるか否かが興味あるところです。

ある女性が子宮頸癌のスクリーニングテストで陽性という結果を受けてあなたの外来を受診しました。“先生、私は癌なのでしょうか？”。これに対してあなたは答えなくてはなりません。あなたはどのように説明しますか？

Bayes' theorem

病気の浸透率を $P(D+)$ 、テストが陽性で病気である確率を $P(D+/T+)$ 、病気でテストが陽性の確率を $P(T+/D+)$ 、病気じゃないのにテストが陽性の確率を $P(T+/D-)$ とします。スクリーニングテストで知りたいのは $P(D+/T+)$ ですが、実際には $P(D+)$ 、 $P(T+/D+)$ 、 $P(T+/D-)$ が判っています。

ベン図を書いてみてください。



$P(T+ \cap D+)$ は全体の中でテストが陽性であり病気である人の割合で、 $P(D+ \cap T+)$ と同じ場所を示しています。しかし、前者は全体に対する割合を示し、後者はテスト陽性者に対する割合です。一見同じものを指しているように見えますが、同じなのは分子であり、分母は異なります。条件付き確率である点に注意してください。ですから $P(T+ \cap D+)$ は、テストが陽性であり病気である人の割合 $P(D+ \cap T+)$ にテストが陽性である確率 $P(T+)$ を掛けてやらなくてはなりません。すなわち $P(T+ \cap D+) = P(T+) P(D+ \cap T+)$ 。逆も真なりで、 $P(T+ \cap D+) = P(D+) P(T+ \cap D+)$ 。この2つをつなぐと $P(D+ \cap T+) = P(D+) P(T+ \cap D+) / P(T+)$ となります。テスト陽性者には病気の場合と病気でない場合があります。すなわち $P(T+) = P(T+ \cap D+) + P(T+ \cap D-)$ すなわち $P(T+) = P(D+)P(T+ \cap D+) + P(D-)P(T+ \cap D-)$ であるのは明らかです。ですから

$$P(D+ \cap T+) = [P(D+) P(T+ \cap D+)] / [P(D+)P(T+ \cap D+) + P(D-)P(T+ \cap D-)]$$

となります。この公式 (Bayes' theorem) の重要な点は最も知りたいテストが陽性で病気である確率を既存のデータから算出できることです。

感度と特異性

感度 (sensitivity) とは病気をテスト陽性として捕まえることができる確率で $P(T+/D+)$ であり、特異性 (specificity) とは病気でないものをテスト陰性として捕らえる確率で $P(T-/D-)$ です。病気の浸透率 $P(D+)$ は既存のデータを用います。 $P(D-) = 1 - P(D+)$ ですし、 $P(T+/D-) = 1 - P(T-/D-)$ すなわちその検査の感度、特異性、病気の浸透度が判っていればテストが陽性で病気である確率を算出できるというわけです。ちなみに疑陽性とは病気でないのにテストが陽性になってしまうことで、 $P(T+/D-)$ であり、偽陽性と特異性の和は 1 になります。一方偽陰性とは病気なのにテストが陰性として

まうことで（いわゆる見落とし） $P(T-/D+)$ 、感度との和は1になります。

子宮頸癌のスクリーニング

子宮頸癌は早期に発見できる疾患の1つです。このPap smear と呼ばれるスクリーニング検査は子宮頸部の細胞を拭ってとってきて、その細胞の悪性度を病理検査技師により判定するものです。しかし検査の陽性、陰性が必ずしも病気を持っている、持っていないこととイコールではありません。アメリカでの大々的な調査の結果、16.25%が偽陰性、すなわち検査は陰性だったが実は病気を持っていた人達でした。悪く言えば“病気を見逃した”こととなりますが、検査にはつきものです。逆に検査が陽性で本当に病気だった人は病気の人の中で83.75%となります。分母が子宮頸癌患者であることを忘れないでください。この検査が陽性で実際病気である確率は先にも触れた通り感度とも呼ばれます。すなわちどれくらいの確からしさでyesをyesと言えるかということになります。逆に検査が陽性であっても実は病気でない、人騒がせなこともあります。さて、このような場合を偽陽性と呼ぶのですが、もしも18.64%がそうであったとしますと、検査が陰性で、実際病気でなかった確率は81.36%となります。これは特異性とも呼ばれます。すなわちNoをNoと言えるかということになります。例えば実際の病気は20%の人にしかないのに全員が陽性になったら感度は100%です。感度100%という優れた検査のように錯覚してしましますが、特異性は0%なのでまったく使いものになりません。一方同じ集団に対して陽性者がいなかったとすると、特異性は100%ですが感度は0%となり、これも駄目検査となります。

特異性、感度ともに高い検査が優れた検査を言えるわけですが、現実的には感度を上げると偽陽性の率もあがり特異性も下がります。逆に特異性を上げようと思うと偽陰性が増えて感度が落ちてしまいます。よって適当な境界線を決めてやらなければなりません。横軸に偽陽性、縦軸に感度を採ったグラフはReceiver operator characteristics (ROC) curve と呼ばれます。直線ではなく対数グラフになるので、適当な落としどころ(cutoff)を決めてやります。

それでは本題に戻ります。Pap smear 陽性で子宮頸癌である条件付き確率はいくらでしょう？

$$P(D1/T+) = P(D+)P(T+/D+) / P(D+)P(T+/D+) + P(D-)P(T+/D-)$$

病気でも検査が陽性である確率は83.75%、病気でなくて検査が陽性（偽陽性）である確率は18.64%なのは解っています。文献を調べて子宮頸癌の一般人口における頻度は0.0083%なので一般人口に占める子宮頸癌でない確率は100%から引けばよく、よって $0.0083 \times 83.75 / 0.0083 \times 83.75 + 99.917 \times 18.64 = 0.0373\%$

となります。すなわち100万人のPap smear 陽性者のうち、本当に癌であったのは373人でしかないのです。ですから“検査が陽性でも病気の可能性が陰性の人よりも4倍位高いので念のため検査をすすめましょう”と陽性者に説明するのが正解です。スクリーニング検査は診断のために行われるわけではなく、次の検査に進むべき人を濃縮するために行う検査であって、陽性率が低いからといって不正確な検査というわけではありません。その点を患者さんに理解してもらう必要があります。かつて子宮頸癌のスクリーニング検査は優れた検査として世界的に受け入れられていましたが、最近ではその有用性について疑問視されています。かつてより偽陰性の頻度が上がったことによるとされていますが、その原因として細胞採取の仕方の不備だけでなく、100万人検査してわずか373人しか子宮頸癌を発見できない精神的、肉体的労力が、発見率を下げてしまっ

たのかもしれませんが。実際細胞診をする人も大変です。

結核スクリーニング

例えば結核のスクリーニング検査で下記の結果を得ました。

x-ray	Tuberculosis (-)	Tuberculosis (+)	Total
Negative	1739	8	1747
Positive	51	22	73
Total	1790	30	1820

結核の頻度は 0.000093 だとします。さて胸部エックス線写真で陽性所見を得た場合、本当に結核である確率は0.239%となります。子宮頸癌より10倍近く高いわけですが、結核の頻度が人口の0.0093%であり(Prior probability)、検査をすることにより25.7倍も結核である確からしさ(Posterior Probability)を挙げたこととなります。

浸透率の計算

先の例では浸透率がわかっていました。今度は逆に浸透率が判っていませんが、スクリーニング陽性者のうち本当に病気であった人の割合に関するデータが判っています。Bayes' theorem は検査だけでなく、ある人口にどれくらい疾患が浸透(Prevalence)を調査するのにも向いています。

Bayes' theorem 公式において検査の陽性率は判っていて病気の浸透率を知りたい場合、公式を少しいじると

$$P(D+) = P(T+) + P(T-/D-) - 1 \quad / \quad P(T+/D+) + P(T-/D-) - 1$$

が得られます。すなわち 病気の浸透率 = (検査を実際に施行した際の陽性率 + 特異性 - 1) ÷ (感度 + 特異性 - 1) で示されます。検査での陽性率は病気の浸透度とイコールではありませんので注意してください。

New York 州マンハッタンで生後 28 ヶ月の小児の血液中の HIV の抗体価を調べました。28 ヶ月の時点で検査を施行した理由は、乳児期では胎盤を介して母親の体内の抗体が移行するためです。検査は匿名で行われ、結果も通知しない形で調査が行われました。結果は 50364 人を対象に調査し、799 人が陽性でした。この調査では母親が HIV positive である場合子供も HIV に感染しているとしました P(H)。検査陽性の中には病気の場合とそうでない偽陽性の場合があります。ここで陽性率 (n+/n) は 0.0159、感度 P(T+/H)=1- P(T-/H) が 0.99 で、特異性 P(T-/Hc)=1- P(T+/Hc) が 0.998 であったとしますと、 $0.0159 + 0.998 - 1 / 0.99 + 0.998 - 1 = 0.0141$ 。そこでマンハッタンにおける 28 ヶ月の小児に HIV が浸透している確率は 14.1% となります。

神経芽細胞腫のスクリーニング

神経芽細胞腫は子供に特異的にみられる悪性度の高い癌です。この癌は血清 NSE や尿中カテコラミン代謝産物の測定が診断の際に有用です。これを利用して生後 6 ヶ月の時

点で尿のカテコラミンを測定して神経芽細胞腫を早期に発見し早期治療を施すことによって予後を改善しようとする試みが成されました。その結果神経芽細胞腫の頻度は約倍に上がったのですが、スクリーニング検査で発見される神経芽細胞腫はほとんどが予後良好のものばかりで、主に1歳以降にスクリーニング以外で発見される予後不良の神経芽細胞腫の頻度を減らすには至りませんでした(Lancet 1996;348:1682)。しかも1歳以降に発見される神経芽細胞腫はスクリーニングでは陰性で、保存尿を再検しても陰性であったのです。以上のことより、生物学的特性の異なる2種類の神経芽細胞腫が存在すると考えられます。1つは良性でおそらくは何も治療しなくても自然に退縮するタイプのもので、スクリーニング検査では主にこのタイプを引っ掛けていたものと思われます。すなわちスクリーニング検査の行われる前は健康な乳児として扱われていたものが、検査によって無用な手術ないし化学療法が行われてしまったわけで、スクリーニング検査が裏目にでたケースと言えます。もう1つは悪性度の高い主に1歳以降に発症するタイプで、おそらくは病気の進行が急速で、スクリーニング以降に発症するものと思われます。日本の1部の施設では18ヶ月でのスクリーニングの効果を検討しています。

ダウン症のスクリーニング

母親の年齢が上がるとダウン症候群の出産の確率が上昇します。胎児がダウン症であるか診断するためには羊水検査、絨毛検査が行われますが、これらの検査によって胎児を流産させてしまう可能性もあり、その適応は慎重に決定されなくてはなりません。特に偽陽性は正常胎児を流産させてしまう可能性を含んでいます。最近妊娠初期〔10 - 13週〕血清中の妊娠関連 protein A と超音波検査、妊娠中期〔14 - 22週〕 α -fetoprotein, 非抱合 estriol, 絨毛 gonadotropin, inhibin A を組み合わせて検査することにより感度は85%、特異性は99.1% (偽陽性0.9%)であり、従来のスクリーニング検査が感度69%、特異性78% - 95%であるのに対して優れた結果が発表されました(NEJM 1999;341:461)。

問題：35歳の妊婦がダウン症児の出産を心配してあななの外来を受診しました。35歳における本スクリーニング方式による感度は92%、特異性97.7%であるとし、この母親の年齢での一般的なダウン症児の産まれる確率は2%とします。あなたはこの妊婦に対してどのような説明をしますか？

解答

1. 障害児を持つ事の意味

産まれてくる子供は誰でも先天異常などの障害をもつ可能性があり、また障害をもって生まれた場合でも様々な成長発達をする可能性がある。すなわち全か無という問題ではないこと。また障害はその子供の個性の一側面でしかなく、障害という側面だけから子供をみるべきではないこと。子供に障害があるからといって家族が不幸になるとは限らない事

2. 検査の対象となる疾患、主にダウン症、神経管欠損に関する医学的説明

3. スクリーニング検査の原理、方法、意味

検査が陽性であった場合ダウン症児を出産する確率は36%であり、逆に64%は違うでしょう。陰性であった場合でもダウン症児を出産する確率は0ではない。

4. その後の選択肢

ダウン症児妊娠の診断のため羊水検査が必要であること。ただし、羊水検査によって1/300の割合で流産が生じること。検査が正常でもダウン症でないとは言いきれない

こと。また他の障害を持つ児を出産しないとは限らないこと。

5 . 必要であれば Second opinion あるいは専門的カウンセリングを受けられるようにすること。

などをプライバシーを保てる状態で口頭説明と文書による補足により判り易く話します。

以上スクリーニングについて実例をもって概説しました。スクリーニングテストは症状が出現する前に病気の細胞や所見を捉えることによって診断しますが、将来的には遺伝子などを調べることにより素因からリスクを割り出すことになるのではないのでしょうか？