HTA における臨床効果評価

はじめに

技術が与える影響を評価するには、医療制度や社会で生じる可能性を反映した包括的な情報が必要になります。優れた分析には、専門家のアドバイスや情報源として用いられるさまざまな学問領域の手法を活用することが求められます。

HTA 評価は比較分析として、既存の標準治療と新しい技術を比較し、新しい技術がどのような価値 (いわゆる「付加価値」) をもたらすのか確認します。国、地域、病院ごとに HTA 評価の実施方法が異なります。国も地域も病院も、その実情に合わせて健康問題を考慮し、規制機関が医薬品に対して認めた指示に関して治療を評価します。

指示の範囲内で、HTA 団体は入手できるデータを調べ、最善の標準治療と比較して治療がどの程度の効果を上げているか (安全性と臨床効果の観点から) 評価します。HTA 団体の一部は医薬品の費用および費用対効果も評価しますが、倫理、組織、社会および法律の側面を正式に評価するところもあれば、こうした問題を自らの評価の中で暗示的に考慮するところもあります。

技術の「付加価値」は各 HTA 組織によって様々に多面的に判断されます。技術の「付加価値」についての結論は、HTA 組織ごとに異なっている場合があります。HTA の欧州ネットワーク (EUnetHTA) は、「付加価値」を評価できる枠組み、 HTA Core Model®

を構築しました。1

HTA Core Model®

には 9 つのドメインがあります。

1. 健康問題
2. 技術の専門的な説明
3. 安全性
4. 臨床的有用性
5. 費用および経済的評価 (費用対効果)
6. 倫理分析
7. 組織的側面
8. 社会的側面
9. 法的側面

EUnetHTA は最初の 4 つのドメインを「相対的効果評価」と定義しました。この評価では、例を挙げると、新しい治療が既存の治療と比較されます。

HTA における臨床効果評価

臨床効果評価では、標準的な臨床環境にいる患者の健康状態に新しい技術が及ぼす効果を調べ、最新の標準治療の臨床環境と比較します。技術が健康に及ぼす影響は通常、健康面での転帰をさらに調べることで分析されます。患者が利用を望む新薬は以下の効能を持ちます。

心臓発作、入院、副作用など「不良」と認識される転帰を削減する、および/または
機能改善や無痛日数など「良好」と認識される転帰を増加させる。

臨床効果評価の際に、HTA 団体は関連した医薬品分野の確立した手法を用います。特に、臨床効果の評価は疫学および医学 (臨床疫学という) の原則を使って実行されます。

良好な臨床効果評価の基本的な原則は 4 つあります。

1. 情報の検索、
2. 妥当な質問の問いかけ、
3. 格差の理解、および
4. 格差の評価。

情報の検索

HTA 団体は臨床情報を使用し、新しい医薬品を投与された際に患者がどのような健康上の転帰になるかについて評価します。とはいえ、まず決定すべきことは情報収集方法です。HTA 団体が新しい技術に関して臨床情報を入手する方法は 3 つあります。

1. 医薬品の効能に関する既存の情報を見直すこと、
2. 新たに試験を実施して、情報を収集し実際の現場で医薬品の効能を評価すること、あるいは
3. 臨床医と患者 (「熟練患者」) に当該医薬品について期待していることを尋ねること。

HTA 団体はよくこうした手法を組み合わせて使用します。例えば:

HTA 団体は当該技術の市販承認取得者 (MAH) の情報を使用して、独自の審査や分析に資する場合があります。
情報がない場合は、短期の転帰 (コレステロール低下など) の変化によって長期の転帰 (入院を避けるなど) の変化
が予測できるかどうか確認するなど、専門家の意見が必要となることもあります。

HTA 団体は新しい試験を委託することはめったにありません。試験を作成して認可してもらうために要する時間が通常あまりにも長くかかってしまうためです。場合によっては、責任団体が追加情報の収集に基づいて、医薬品の条件付き償還を認めてきました (これは規制当局が、追加情報の収集が必要となる条件付市販承認 (MA) を認可することに類似しています)。新薬が実際に期待通りの働きをしない場合のリスクについては、価格交渉メカニズムによって、あるいは償還手続き関連の条件に対する他の変更 (償還手続きを受け取る資格がある患者集団に対する追加規制など) によって MAH と責任団体の間で共有することがありますが、患者はより直接的なアクセスを提供されます。

妥当な質問の問いかけ

新たな医療技術の臨床効果を評価する際は、HTA 団体はその技術および関連するすべての転帰を注意して考慮しなくてはなりません。技術の効果について妥当な質問の問いかけができるように、これらの転帰について知ることが重要です。

臨床医にとって重要に思える転帰が、患者によって最も重要なものとみなされるとは限りません。このような理由で、患者が試験の設計に関与し、患者にとって重要な転帰に関する情報の収集を確保することが重要です。たとえば、近年認識されてきたことですが、生活の質は患者の重要な転帰のひとつです。このため、生活の質の尺度、および臨床試験におけるいわゆる「患者による転帰評価」を作成する特定の方法が開発されています。

特定の技術についてすべての重要な転帰を調べる手法の 1 つは、分析的フレームワーク、たとえば図 1 にあるようなフローチャー

トを使用することです。2

分析的フレームワークは介入に関連する転帰のすべてを視覚化したり、不確実性がどこにあるのか強調表示した
りする役に立ちます。

図 1 の分析的フレームワークにおいて

原因と結果は矢印で示されています。
曲線の矢印は有害な転帰を示します。
健康状態改善の転帰 (死亡率の低下など) は矩形で示されています。
鋭角の矩形は臨床的意義のある評価項目 (胸痛など患者が認識する評価項目) を示します。
丸角の矩形は代替評価項目 (血中コレステロール値など患者が認識できないもの) などの中間評価項目を示
します。
不確実性に関連する主要な疑問を数値的に示すことができます。

この分析フレームワークは、小児および思春期のディスリピダミア（脂質代謝の障害）のスクリーニングの有効性に関するエビデンスの利点と制限を明らかにするために使用されました。ディスリピダミアは冠状動脈性心臓病（CHD）の重要なリスク要因です。(免責事項：この画像は、信頼性の高いAI支援翻訳ツールを使用して翻訳されました。その精度と多言語の幅広い能力が証明されています。)

この分析的フレームワークに関連する主要な疑問は次のようなものです。

主要な疑問 1:小児および青年の脂質異常症のスクリーニングは、発症を遅延し CHD (冠動脈性心疾患) に関連し
た事象の発生を低減する際に有効でしょうか。
主要な疑問 2:小児および青年が CHD 関連の事象やその他の転帰のリスクが高いことを特定する際に、脂質異常症の
スクリーニングの正確性はどれくらいあるでしょうか。
主要な疑問 3:スクリーニングの副作用はどのようなものでしょうか (偽陽性、偽陰性、ラベリングなど)。
主要な疑問 4:小児および青年において、成人期の脂質異常症の発生率を低減し、CHD 関連の事象や他の転帰 (治療開始の最適年齢など) の発現を遅延させ、またその発生率を低減する際に、薬物、食事、運動および併用治療の効果はどれくらいあるでしょうか。
主要な疑問 5-8:小児・青年の脂質異常症の治療に関して、薬物、食事、運動および併用治療の効果 (小児期に脂質異常症を治療する増分利益も含めて) はどれくらいあるでしょうか。
主要な疑問 9:小児および青年において、医薬品、食事、運動および併用治療の副作用はどのようなものでしょうか。
主要な疑問 10:小児期に脂質異常症を改善することで成人期の脂質異常症のリスクは減るでしょうか。
主要な疑問 11 (写真なし):無症状の小児における脂質異常症のスクリーニングに関連する費用の問題はどのようなも
のでしょうか。

転帰の格差の理解

すべての重要な転帰が特定されたとしても、新しい技術の効果を標準治療や他の既存の治療と比較する際に依然としていくつかの課題が存在することがあります。転帰は異なった方法で評価される場合もあり、2 つの技術がより詳細な検査で違いが示されるまでは類似した転帰を持っているように見える場合もあります。

重要で特定された転帰が評価しにくい場合や、以前にまったく評価されなかった場合には、科学者が臨床試験で再現可能な評価基準を注意深く策定しなければなりません。たとえば、ある患者が医薬品によって自分がどのようにしたら職場に復帰できたり、ベッドから出ることができたりするのか知りたいと思っているとします。科学者は腰痛を抱える患者のために数値による疼痛評価スケールを作成するかもしれません。他のケースでは、たとえば臨床試験で検査項目の変化を評価した場合には、この変化は職場に復帰できるかどうかなど、患者にとってより重要な評価基準で再解釈する必要があります。

時には、医薬品の製造業者が血圧の低下などの短期的な転帰によって新しい医薬品の効果を示したことに、医薬品を承認する規制当局が満足するかもしれません。HTA 団体は、短期的な転帰について、早期の死亡回避など患者に一層関連した転帰へと再解釈する必要があります。

転帰の中には直感的に見えるものもありますが、よく調べてみると、解釈しにくいものもあります。たとえば、5 年死亡率 (5 年以内に死亡) のリスクが 50% 低下しても、医薬品によって早世が予防できるわけではありません。簡単に言えば、次のようになりま

す。

一部の患者で平均余命が 4.9 年から 5.1 年に延びる (またはさほど改善せず 4.99 年から 5.01 年に延びる)、あるいは
ごく一部の患者における疾患の治癒があるものの、他の患者においては全く平均余命は延びない。

患者にとって有意な評価項目に格差が認められたとしても、こうした格差はやはり解釈しづらい場合があります。たとえば、臨床試験で、新しい医薬品が感染による入院のリスクを 33% 低減したと指摘されるとしましょう。しかし、これは異なった意味合いを持つことがあります。たとえば、以下のような意味です。

薬を飲んだ100 人のうち 33 人が、飲まなかったら入院した可能性がありましたが、実際には入院しないで済みまし
た (これは、リスクの絶対的軽減となります)。または
服用なしに入院する可能性と比較すると、入院の可能性が 33% 低下しました (これは、リスクの相対的軽減となりま
す。新薬がない場合に入院する可能性を 1,000 人に 3 人とすると、33% の低下がある場合、入院の可能性は減り、1,000 人に 2 人となります。言い換えると、薬を飲んでいた 1,000 人に 1 人がその恩恵を受けます。これは、100 人のうち 33 人が恩恵と受けた上述の例とは全く異なっています。

新しい医療技術と標準治療の格差を理解するうえで最終的な課題となるのは、統計的検査の利用と誤用です。統計的検査の意図は、検出した格差が実際の格差となりうるか研究者が把握できるようにすることです。大抵、これは p 値という形で報告されます。しかしながら、 p 値は格差の程度 (大きさ) や、患者にとって格差が有意であるかどうかを反映していません。つまり、p 値は通常は決定を下す患者や提供者にとって役に立つものではありません。

他の統計上の評価項目には信頼区間も含まれています。信頼区間がより有用となるのは、それによって新しい医療技術と標準治療の格差の大きさに相当の意味が与えられるからです。信頼区間は、格差の程度の推測について不確実性も反映しています。たとえば、新しい医薬品が、心臓発作を起こす現在の可能性と比較すると、心臓発作が今後起こる可能性を 33% 低下させる (信頼係数 95% で 5% から45% の信頼区間) と報告されるかもしれません。

格差の評価

最後の課題は、転帰の格差を認識し評価する方法を理解することです。ある医薬品によって、寿命が 0.2 年延びたとしても、HTA 団体はやはり次のことを知る必要があります。

予測された副作用や他の懸念を考慮して、患者が 0.2 年の余命延長をどの程度評価するか、
すべての患者がほぼ同じような余命延長を経験するかどうか、または患者に劇的な格差があるかどうか、およびすべての患者がこうした余命延長を同じように評価しているか。

平均余命が 0.2 年延びた新しい医薬品に関しては、格差がほとんどなくすべての患者で 0.2 年延びた場合と比較すると、一部の患者には効果があったものの他の患者では効果がなかった場合では、受け止め方が異なります。

健康転帰の格差に対して患者や医療提供者が与える相対的価値を理解するために利用できる機序がいくつかあります。1 つはフォーカス

グループの調査などの定性調査で、どの転帰が患者にとって最も重要であるのかを理解することを意図したものです。もう 1 つは患者調査に基づいた定量調査で、健康状態の違いを示す重要度を正確な数値で示そうとするものです。

要するに、臨床効果の評価として、以下の質問に対処することが推奨されます。

1. 情報はどの程度包括的なのか。
2. 情報はどの程度正確なのか。
3. 何か欠けていないか。
4. 情報はどの程度理解できるか。

参照文献

1. HTA Core Model.Retrieved 7 December, 2015, from http://www.eunethta.eu/hta-core-model
2. U.S. Preventive Services Task Force (2015).Final Update Summary:Lipid Disorders in Children:Screening.Retrieved
7 December, 2015, from:http://www.uspreventiveservicestaskforce.org/Page/D ocument/UpdateSummaryFinal/lipid-disorders-in-children-screening

A2-6.03.1-v1.1