

NECA - 기본연구

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

2009. 12.

연구 경과

연구시작일

2009년 5월 1일

연구종료일

2009년 12월 31일

기획위원회 심의일

2010년 1월 18일

연구성과 검토위원회 심의일

2010년 4월 2일

최종보고서 수정완료일

2010년 6월 1일

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

주의

1. 이 보고서는 한국보건 의료연구원에서 수행한 연구사업의 결과보고서입니다.
2. 이 보고서 내용을 인용할 때에는 반드시 한국보건 의료연구원에서 시행한 연구사업의 결과임을 밝혀야 합니다.

연구진

연구책임자

이상무

한국보건의료연구원, 연구위원

참여연구원

김수영

한림대학교 의과대학, 교수

신승수

아주대학교 의과대학, 교수

서현주

한국보건의료연구원, 의료기술분석팀 책임연구원

장보형

한국보건의료연구원, 의료기술분석팀 책임연구원

박주연

한국보건의료연구원, 경제성분석팀 연구보조원

자문위원

장은진

한국보건의료연구원, 보건의료성과분석팀 책임연구원

목차

Executive Summary	i
요약문	v
1. 서론	1
1.1. 연구배경	2
1.2. 연구목적	5
2. 연구방법	6
2.1. 치료효과 크기 지표 분석	7
2.2. 효과크기에 따른 의사결정 방법론	7
2.3. 1개 질환 치료효과크기 mapping	7
2.3.1. 질환 선정 기준	7
2.3.2. 질환 후보	8
2.3.3. 치료효과 크기의 mapping 방법	8
3. 연구 결과	10
3.1. 치료효과 크기 지표 분석	11
3.1.1. 치료효과 크기 지표 개념	11
3.1.2. 치료효과크기 지표에 따라 의사결정이 달라진 사례연구	34
3.2. 효과크기에 따른 의사결정 방법론	37
3.2.1. 여러 근거 중심 의사 결정에서 BRA	37
3.2.2. 여러 기구에서 BRA	49
3.2.3. 주요 위험-이득분석	52
3.3. 급성 관상동맥 증후군의 표준치료법에 대한 NNT mapping	71
3.3.1. 초기 약제 치료(Early pharmacological intervention)	71
3.3.2. 초기 질병 관리(Initial management)	76
3.3.3. ST 상승 급성관상동맥증후군에서 재관류요법(12시간 이내의 일차 PCI 와 혈전용해에 초점을 둠)	79

3.3.4. 침습적인 검사와 재관류요법	80
4. 고찰 및 결론	81
4.1. 고찰	82
4.2. 결론	84
5. 참고문헌	85
6. 부록	94
부록 1. The 2007 Oxford league table of analgesic efficacy(Oxford 2007)	95
부록 2. Acute Coronary Syndrome의 치료효과크기 자료추출 결과	96
부록 3. 근거중심 의사결정 방법론 심포지엄	108

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

표 차례

표 1. 5개 후보질환에 대한 최신 업데이트된 guideline 검색	8
표 2. NNT 전환식	9
표 3. 새로운 가상의 신약의 수술 후 합병증 발병에 미치는 영향에 관한 가정 예	13
표 4. 가상의 두 치료법에 대한 여러가지 효과 측정방법에 따른 결과값	19
표 5. Intervention의 효과를 표현하는 다음 5개의 효과크기: odds ratio, relative risk, risk difference, weighted mean difference, standardized mean difference ¹⁾	32
표 6. 두 의료기술에 대한 급의사결정에서 임상연구 결과 제시방법에 따른 정책결정자들의 판단. 평균 점수(%) (95% CI) (Fahey 1995)	35
표 7. 여러 다른 지표들로 제시된 사람들의 스타틴 복용 의사결정 (Carling 2009)	35
표 8. 원하는 결과와 원하지 않는 결과 사이의 저울질	47
표 9. 류마티스 관절염 치료 약제간의 RV NNT	62

그림 차례

그림 1. 대조군에서의 결과발생의 빈도(P_0)에 따른 RR과 OR과의 관계	14
그림 2. 중등도 이상의 통증에서 복용 후 4-6시간 지나서 동통의 50%가 감소되는 것을 효과로 보았을 때 NNT값의 대진표	20
그림 3. 임상시험과 체계적문헌고찰에서 얻어진 NNT와 NNH	21
그림 4. ARR 및 RRR값으로 표현한 경우 효과에 대한 인식 정도	34
그림 5. 절대위험도감소 대비 상대위험도감소의 효과에 대한 판단의 비교	36
그림 6. 아스피린의 대장암 사망률 감소 효과	40
그림 7. 아스피린의 대장암 발생률 감소	40
그림 8. 아스피린의 대장 선종 발생 감소	41
그림 9. 선종 환자에서 아스피린의 선종 예방 효과	41
그림 10. 아스피린의 심혈관 관련 유해	42
그림 11. NSAIDs의 대장암 사망률 감소 효과	43
그림 12. NSAIDs의 대장암 발생 감소 효과	43
그림 13. NSAIDs의 선종발생 감소 효과	44
그림 14. NSAIDs의 심혈관 질환 유해	44
그림 15. NNT threshold, target event에 대한 환자의 기저 위험도, NNT간의 관계	57
그림 16. value tree	69
그림 17. ACS Early Pharmacological intervention NNT mapping	72
그림 18. Early Pharmacological intervention의 NNT mapping을 Lowess 방법으로 smoothing 한 결과	73
그림 19. ACS Anticoagulant therapy의 NNT mapping	74
그림 20. ACS Statin therapy의 NNT mapping	74
그림 21. ACS Beta Blocker therapy & Antianginal therapy의 NNT mapping	75
그림 22. ACS Angiotensin Receptor Blockers therapy NNT mapping	75
그림 23. ACS Initial management NNT mapping	76
그림 24. Non-ST elevation ACS의 Initial management NNT mapping	77
그림 25. ST elevation ACS의 Initial management NNT mapping	77
그림 26. ACS Antiplatelet therapy NNT mapping	78
그림 27. ACS Anticoagulant therapy NNT mapping	78
그림 28. ACS Beta Blocker therapy NNT mapping	79
그림 29. ST 상승 ACS 환자에서 Thrombolysis에 비교된 Primary PCI의 NNT mapping	79

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

그림 차례

그림 30. Primary PCI로 치료받은 ST 상승 ACS환자에서 placebo와 비교된 Glycoprotein IIb/IIIa receptor antagonists(abciximab)의 NNT mapping	80
그림 31. non-ST 상승 ACS환자에서 conservative approach와 비교된 routine 관상동맥혈관성형술과 재관류요법	80

Executive Summary

Background

The rapid turnover in the means of registering new drugs and new health technology assessment has been led by the increase of interest in evidence-based decision making as well as its effects on the clinical guidelines.

In these days, evidence-based decision making in healthcare begins to be increased by regulations of new health technology assessment, registration of new drugs and interest in the clinical practice guideline.

The process for evidence-based decision making involves the process gathering the information about benefit and harm of intervention through systematic search, and assessing quality of individual studies. Based on these results, the policy makers and clinician could make a decision based on the evidence based-information.

The most important part of this process is to weigh up the beneficial and harmful effects of intervention, which is weighing the benefits and risks to determine the trade off between the two.

However, the decision making process may be affected by how effect size is presented and analyzed even for a same study.

Also, there is an increasing need of benefit risk analysis in offices that determines the benefit coverage and approval of various drugs.

Therefore, we aimed to investigate the how decision makers are recognize the intervention according to presentation of various effect size, and to examine the current status of benefit risk analysis which allow to incorporate policy makers' perspectives as well as patients' one.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Methods

We performed the comprehensive searches using evidence based medicine textbooks, Cochrane handbook and various related journals. Also, offices involved with sanctioning new drugs such as FDA, CHMP (Committee for Medical Products for Human Use) have been searched for further information.

To inquire the acceptable NNT range of standard and traditional therapies, we selected the diseases with acute/chronic progression and with outcomes of interest that include from symptom alleviation to death. Also, the selected disease is restricted to a specific organ without externalities, and has an international standardized treatment guidelines.

The relationship between the magnitude of effects and utility has been represented in a two-dimensional graph (Y axis: utility, X axis: NNT)

Results

There are advantages and limitations in measuring the treatment effects of interventions. Relative effect size such as Relative risks (RR) and Relative Risk Reduction (RRR) is useful indices that may be summarized and analyzed easily from clinical, cohort and randomized controlled studies. However, these cannot be used in cross-sectional and case-control studies. Also, since it relies on the outcomes' probabilities, their use is limited in randomized controlled trials or in selecting the study participants by representative sampling. Also, since it is unable to differentiate between very small effects in absolute number and remarkable treatment effects, it may be risky to provide only relative treatment effect size with nature of relative comparison. Therefore, it is needed to provide absolute risk reduction values or NNT(Number needed to treat) values to provide values that may be utilized for absolute comparisons. Especially, the NNT is the

number of patients who need to be treated in order to prevent one additional undesirable outcome (i.e. the number of patients that need to be treated for one to benefit compared with a control in a clinical trial). Also, it provides comparative information between studies/treatment when measured from comparison groups with similar underlying risks during same periods.

However, only few studies provide NNT or ARR values in representing treatment effect size. Also, since medical researchers as well as decision makers lack in knowledge about NNT values, transfer of knowledge and implementation of the education about absolute effect size measures in this field is needed.

However, there are some limitations to NNT value. Since it is affected by variables such as magnitude of underlying risk, type of comparison group and the period when the clinical data was obtained. Even though it may be effective in comparing in between different treatment effects, these variables must be taken into consideration during the comparison process. Because NNT does not incorporate values of patients or society in itself, variables such as utility, cost, and adverse effects must be taken into consideration during the decision making process.

Also, there is a need to pay an attention to the results of outcomes representing by various effect size measures may alter the interpretation of treatment effects. Some studies have showed that decision makers might be overestimated the treatment effect in the relative effect size measures than the absolute ones.

In deciding health policy affecting the whole nation, decision makers are usually followed by criticisms as well as controversial debates. Therefore, there has been a significant need of utilizing objective methods for balancing the benefits and risks of various treatment methods. However, despite the efforts to find an objective method, an international agreement in choosing the best evaluation method has not be made yet. Regardless of a unanimous agreement on the

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

evaluation method, benefit-harm tradeoff must be a transparent and objective process in the decision making.

As proposed by CHMP, an ideal method allows the user to clearly identify the benefits and risk, and decide weight on important and less important risks/benefits while providing information about evidence strength and identified uncertainties. However, it might not be applicable in real world that have to deal with many variables.

According to US FDA's requirements of new benefit risk analysis, the method must be able to capture the complexities of benefit- risk analysis, quantify the nature of benefit/risks in individual patients, and to be updated with new information. In addition, it needs to be able to deal with various uncertainties while having to deal with cost issues in benefit risk analysis. It should be able to provide and share quantitative information while dealing with cost issues and potential adverse effects.

There is a need to develop a comprehensive analysis method that is easy to perform a trade off between benefits and risks in every-day practice.

Conclusion, Proposal and Consideration

In making a decision based on evidence, it needs a sufficient amount of data obtained from systematic reviews on existing studies. The results obtained from various studies have its advantages and limitations.

Findings of studies using binary outcome data should be provided both relative and absolute effect measures to help interpret. Also, the decision makers should be able to understand various features of provided treatment effect size.

Since various trade-off methods of risks and benefits are derived from previous studies, the decision makers must understand the advantages and limitations of each method, and then utilize them when making a decision whether to choose a specific treatment.

요약문

연구배경

최근 국내에서도 신의료기술의 평가, 신약의 등재과정 등 보건의료분야에 근거중심의사결정 체계가 급속도로 제도화되기 시작하였고 임상진료지침에 대한 관심도 증가하고 있다. 근거중심의사결정의 과정은 기존의 연구 결과에 대한 체계적인 검색과정과 문헌에 대한 비평적 평가를 통해 얻는 해당 중재의 효과와 위해에 대한 정보이며 이를 바탕으로 해당 중재를 사용할지에 대한 의사 결정을 하게 된다.

이 과정은 특정 중재의 이득(benefit)과 위험(risk)을 평가하고 둘 사이의 저울질(trade-off)을 하는 최종 과정이 필요한데 이러한 연구를 통해 얻어진 치료의 결과를 어떻게 표현하느냐에 따라 해당 치료의 이득에 대한 치료적 효과의 인식이 영향을 받게 되며 결과측정 방식에 따라 동일한 치료법에 대해서도 효과의 크기에 대한 인식이 달라질 수도 있다.

또한 약제의 인허가 관청이나 급여 여부를 결정하는 기관에서도 이러한 이득-위험 분석의 필요를 느낄 수 있다.

이와 같은 배경 하에 치료법에 대한 잘 고안된 연구를 통해 얻어진 치료의 효과 측정법에 대한 올바른 이해와 그 치료법들이 어떻게 인식되는지 연구할 필요가 있고 이러한 이득과 해로움에 대한 가치 판단과정에 있어서 보다 더 객관화 시킬 수 있는 방법은 없는지 알아 볼 필요가 있다.

연구방법

치료효과를 측정하는 방법에 대한 분석과 효과크기에 따른 의사결정 방법론에 대한 고찰은 근거중심의학 관련 교과서, Cochrane handbook, 관련 학술지에 대한 고찰을 통하여 이루어졌으며 FDA, CHMP(Committee for Medicinal Products for Human Use) 등 신약의 승인과 관련된 기관을 검색하여 관련 방법론을 고찰하였다.

기존 치료법들에 대한 근거중심임상진료지침을 이용하여 의학적으로 받아들여진 치료법에 대한 효과의 크기와 그 측정값의 효용을 도식화 하여 의사결정의 참고수치화 할 수 있는 여부를 알아보기 위한 연구를 위해서는 질병의 자연사가 급·만성적인 경과과정이 있고, 의료결과가 증상완화에서 사망까지를 포함한 질환으로 외부효과가 없는 특정 기관(organ)에 관련된 질환으로서 국내외 임상진료지침에서 제시한 표준적인 치료법을 갖고 있는 질환을 선정하였다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

효과의 크기와 효용의 관계성을 2차원 그래프(y축: utility, x축: NNT)로 도식화 하였다.

연구결과

치료의 효과를 측정하는 방법에 관해서는 측정법의 특성에 따라 장점과 한계점들을 갖고 있다. 비교위험도(RR)이나 상대위험도감소(RRR)같은 지표들은 임상시험의 결과들을 요약 하는데 유용한 지표들이며 비교적 이해하기 쉽고, 매우 다른 임상적 상태에서도 RRR값은 같을 수 있으며 전향적 코호트 연구나 무작위대조군 임상시험에서 구할 수 있으나 그 외 단면 연구나 환자-대조군연구 같은 경우 사용할 수 없다. 결과의 확률에 기초한 것이므로 무작위 선정이나 대표성이 있는 표본이 선정된 경우에 한하여 사용할 수 있다는 제한점을 지니고 있다. 또한 현저한 치료의 효과와 절대수에 있어 매우 적은 효과를 분별하지 못할 수 있으므로 이러한 상대적 비교 특성을 지닌 결과 값만 제시하는 것은 위험할 수 있다. 따라서 이러한 점을 보완할 수 있도록 절대적 비교의 특성을 갖는 절대위험도감소나 NNT 같은 지표를 제시할 필요가 있고 특히 NNT는 한 바람직하지 않은 사건(事件; event)을 막기 위해 얼마의 노력을 들여야 할지를 보다 더 명확히 표현하며 동일비교 대상, 동일 위험군에서 유사한 시기의 결과값을 측정할 경우 연구간, 치료법간 비교에 정보를 제공하게 된다. 그럼에도 불구하고 실제로 소수의 문헌에서만 NNT 혹은 ARR을 보고함으로써 불균형적인 보고를 하고 있다. 실제 우리나라에서도 NNT에 대한 인식이 아직 부족하고 그 개념을 이해하는 의사결정자들이 부족하므로 이 분야에 대한 교육과 지식의 확산이 절실하다.

NNT가 지니고 있는 한계점도 또한 분명하여 기저 위험의 정도, 대조군의 종류, 임상연구에서 결과가 측정된 시기 등에 영향을 받게 되므로 치료간의 효과를 비교하는데 도움이 되지만 이러한 요소를 고려하지 않고 비교하는 오류를 범하지 말아야 한다. NNT수치 자체는 다른 지표와 마찬가지로 그 결과값에 대한 가치가 반영되지 않은 수치이므로 NNT값만 갖고 판단할 수는 없으며 그 효용과 치료에 드는 비용 그리고 치료에 의한 부작용에 대한 고려를 해야 올바른 판단을 할 수 있다.

또한 각 지표가 갖는 특성자체도 있지만 어떤 지표를 제시하는가에 따라 치료의 효과를 인식하는 정도도 달라질 수 있으므로 주의를 요한다. 많은 연구들에 있어서 정도의 차이가 있으며 상당히 이질적이긴 하나 일관되게 상대적 특성을 갖는 결과 측정법에 의해 의사결정하는 것이 절대적 특성을 지닌 결과값 측정법에 의한 정보를 제시한 경우보다 효과에 대해 더 크게 인식된다는 연구에 대해 주의할 필요가 있다.

국가적으로 중요한 의사결정을 할 경우 그 의사결정의 결과에 대해 후에 비평을 받게 되는 경우가 종종 생기게 되며 약제의 승인과정에 있어서 인허가의 의사결정에 대한 시비가

클 수 있다. 따라서 이런 의사결정을 하는 조직이나 위원회는 보다 객관적으로 치료법의 득과 실을 판단하는 방법이 필요하며 이에 대하여는 많은 연구들이 이루어져 왔으나 아직 어떤 방법이 최선의 방법인지 국제적으로 합의된 부분은 부족하다. 하지만 득과 실에 대한 trade off는 의사결정이나 권고사항 결정과정에서 피할 수 없는 부분이며 다양한 이해당사자들이 존재하는 보건의료분야의 특성상 투명하고 객관화 시키는 노력을 기울이지 않을 수 없다. 이상적인 방법은 CHMP에서 제시한 바와 같이 가장 중요한 이득과 가장 심각한 위험을 쉽게 확인하게 해주는 것, 개별 이득과 위험에 대한 명백한 가중치를 부여하는 것, 근거의 강도와 확인된 불확실성에 대해 제시하는 것이며 이러한 원칙하에 각 치료법이 갖는 부차적인 이득과 위해를 종합적으로 반영하고 총합하여 결론을 내릴 수 있다면 좋겠지만 많은 변수를 고려할수록 방법론이 복잡해지고 실제로 사용하기에 어려울 수도 있다.

미국 FDA가 새로운 이득-위험 분석의 조건으로 제시한 바에 따르면 약제 이득 위험 분석의 복잡성을 잘 포착하여야 하며 개별 환자에서 이득-비용의 특성을 양적으로 표현할 수 있어야 하고 새로운 정보가 나올 때 갱신할 수 있어야 한다. 또한 이득-위험 평가에 내재한 불확실성을 다루어야 하고 이득-위험 평가에서 비용의 문제를 해결할 수 있어야 한다. 양적 체계 도입의 비용 문제와 잠재적 이상 효과에 대하여 다루어야 하고 양적인 정보를 효과적으로 제시하고 교류할 수 있게 하여야 한다.

지금까지 연구를 토대로 향후 일상적으로 사용하기에 어렵지 않으면서 총괄적으로 이해득실을 평가할 수 있는 방법론의 도출이 필요하다.

결론 및 정책제언, 고찰 및 결론

근거중심의사결정에 있어서 기존의 임상연구들에 대한 체계적이고 비평적인 방법을 통해 합성된 양질의 근거가 필요하며 이러한 연구들에 의하여 제시되는 결과의 측정 방법들은 각각 특징점과 제한점을 가지고 있다.

치료중재법의 치료효과 크기를 고려한 의사결정에서 효과의 크기를 비교 치료법에 대하여 상대적으로 표현하는 방법들과 차이를 절대적으로 표현하는 방식에 따라서 효과의 크기에 대한 인식이 달라질 수 있으므로 연구자는 결과값을 제시할 때 상대적 특성을 갖는 방법(RR, RRR)과 절대적 특성을 갖는 방법(ARR, NNT)값을 동시에 제시할 것을 권한다.

치료의 이득과 손실을 저울질하는데 있어 보다 객관화된 방법들이 도입되고 있다. 상황에 적절한 치료법에 대한 이득과 해로움을 측정하는 방법의 특성을 사용하여 의사결정과정에서 보다 객관화를 기할 필요가 있다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

1. 서론

1.1 연구 배경

보건의료분야에 근거중심의사결정 체계가 우리나라에서도 2007년 이후 급속도로 제도화되기 시작하여 의료법상 신의료기술평가를 거쳐 약제를 제외한 신의료기술에 대해 근거에 입각한 평가가 국가적으로 이루어지게 되었고 신약의 등재 시에도 임상적 효과에 대한 자료와 더불어 경제성 평가자료의 제출을 요구하여 급여여부의 의사결정에 사용하게 되었다. 또한 기등재 의약품에 대해서도 일정 기간에 걸쳐 재평가를 시행하여 목록을 정비하는 과정에서 임상적 유용성에 대한 근거자료를 분석하고 필요시 경제성 분석을 수행하여 의사결정에 반영하도록 하고 있다. 또한 심사평가원에서 급여 기준에 대한 객관성과 수용성의 증가를 위해 근거에 따른 급여기준을 만드는 노력을 기울이고 있고 의료계에서도 임상진료지침의 제정에 대한 관심이 증가하고 있다.

근거중심의사결정의 근간은 근거중심의학의 개념의 도입됨으로써 이루어진 것이며 근거중심의학은 ‘환자의 문제에 대해 결정을 내릴 때 세심하고 주의 깊게 최신의 의학지식을 적용하는 것이며, 개인의 임상경험과 체계화된 연구에서 얻어진 임상적인 근거들 중에서 최선의 것을 통합하여 개개인의 환자에 적용하는 것’이라고 정의할 수 있다(Smith 1997).

근거중심의학의 영역이 확대되면서 근거중심보건의료의 개념이 나오기 시작하였으며 체계적 문헌고찰, 임상진료지침 등 연구 영역이 확대되고 의료기술평가 등 정책 등에도 근거중심의학의 원리가 많이 이용되기 시작하였는데 연구 결과에 대한 체계적인 검색, 평가를 통해 얻는 것은 해당 중재의 효과, 효능, 효율(이득), 부작용, 비용, 부담(위험)에 대한 정보이며 이를 바탕으로 해당 중재를 사용할지에 대한 의사 결정을 하게 된다.

결국 근거를 적용한다는 것은 특정 중재의 이득(benefit)과 위험(risk)을 평가하고 둘 사이의 저울질(trade-off)을 하는 것이다. 이러한 저울질을 이득-위험 분석(benefit risk analysis)라고 한다(Holden 2003).

이득은 결국 사망률 감소나 입원을 감소, 삶의 질 향상 등과 같은 원하는 결과라고 할 수 있다.

위험은 치료와 관련된 나쁜 것으로 결국 원하지 않는 결과라고 할 수 있다. 위험은 치료로 인하여 발생하는 이상 사건(adverse events)과 치료로 인해 발생하는 불편(약제 복용, 병원 방문 등), 비용 모두를 말한다. adverse event를 safety와 tolerability로 구분하기도 하는데 전자는 실제로 혹은 인지된 치료와 관련된 신체적, 정서적 위협이라고 할 수 있으며 후자는 safety 수준까지 도달하지 못한 위협이라고 할 수 있다(Felli 2009).

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

이득-위험 분석은 개인 차원에서 하는 것과 기관 차원에서 하는 것으로 구분할 수 있다. 전자는 환자에게 특정 치료를 계획하는 의료인이 특정 치료의 이득과 위험을 저울질하여 해당 치료를 할 것인지를 결정하는 것이다. 이때 판단은 정성적으로 할 수도 있고 이득과 위험을 정량적으로 판단하여 결정할 수도 있다.

이러한 연구를 통해 얻어진 치료의 결과를 어떻게 표현하느냐에 따라 해당 치료의 이득에 대한 치료적 효과의 인식이 결정되는데 결과측정 방식에 따라 동일한 치료법에 대해서도 효과의 크기에 대한 인식이 달라질 수 있다.

유방 촬영술과 심장 재활치료에 대하여 임상시험의 결과 효과 크기에 대한 정보를 상대 위험도감소(RRT:relative risk reduction), 절대위험도 감소(ARR:absolute risk reduction), 무사건환자의 비율(proportion of event free patients), NNT(number of patients needed to be treated)로 각각 나타내어 13개 보건 당국의 182명의 위원회 의사결정자들에게 급여의 의사를 질문한 결과 각 효과의 크기를 나타내는 지표에 따라 의사결정이 달라졌으며 가장 높은 점수를 받은 지표가 상대위험도감소였다(Fahey 1995). 그 외 여러 문헌에서도 일관성 있게 상대위험도 표현이 절대적 위험도 표현보다 효과의 크기를 더 크게 인식하는 것으로 나타났으며 각 치료의 측정법이 갖는 의학 통계학적 의미 외에도 인식의 측면에서 어떤 영향을 주는지에 대한 올바른 이해가 있지 않다면 이러한 과학적 근거를 가지고 의사결정하는 과정에 있어서 오류를 가져올 수 있다.

또한 FDA나 우리나라 식약청처럼 허가 관청이나 급여 여부를 결정하는 기관에서도 이러한 이득-위험 분석은 필요하며 일부 기관에서는 해당 제약회사나 의료기기 회사에 필요한 이득-위험 분석을 요청하기도 한다.

최근까지도 이러한 규제 기관에서 어떻게 이득-위험 분석을 하는지 잘 알려져 있지 않았다. 대부분 위원회 위원들의 주관적 판단에 의하거나 어떤 기준으로 판단을 하더라도 정량적인 방법보다는 정성적인 방법을 더 많이 사용하였다.

하지만 몇 가지 사건을 거치면서 이러한 관행에 변화가 요청되게 되었다. 대표적인 사건으로 머크사에서 개발한 Vioxx가 심장마비와 뇌졸중 위험을 높일 수 있다는 우려가 있어서 전세계 시장에서 회수 조치된 것이다. 머크사는 바이옥스가 관절염 치료 외에 대장폴립 재발을 차단하는 효과가 있는지를 확인하기 위한 3년간의 임상시험 결과 바이옥스를 18개월 이상 복용한 경우 심장마비와 뇌졸중 위험이 2배 증가한다는 사실이 밝혀져 임상 시험을 중단하고 한국을 비롯한 전세계 시장에서 이 약물을 회수하기로 한 것이다.

우리나라에서도 유사한 사례가 몇 차례 있었다. 미국 FDA에서 PPA 성분이 출혈성 뇌졸중의 위험성을 높일 가능성이 있기 때문에 2000년 11월 사용하지 않도록 강력히 권고

하였지만 우리나라의 경우는 2002년 연구팀을 구성하여 추가적인 연구를 시행한 이후에 2004년 8월에야 해당 약제사용을 중단시켰다.

이와 같은 배경 하에 치료법에 대한 잘 고안된 연구를 통해 얻어진 치료의 효과 측정법에 대한 올바른 이해와 그 치료법들이 어떻게 인식되는지 연구할 필요가 있고 이러한 이득과 해로움에 대한 가치 판단과정에 있어서 보다 더 객관화시킬 수 있는 방법은 없는지 알아볼 필요가 있다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

1.2 연구 목적

본 연구는 치료법에 대한 치료의 효과 측정법에 대한 올바른 이해와 그 치료법들이 어떻게 인식되는지 또한 이러한 이득과 해로움에 대한 가치 판단과정에 있어서 보다 더 객관화시킬 수 있는 방법은 없는지 알아보기 위하여 효과 크기를 나타내는 지표에 대한 개념을 조사하고, 효과 크기를 나타내는 지표를 활용한 trade-off의 의사결정관련 연구방법론을 고찰하여하여 각 효과크기 해석 시 고려할 사항 제시하고자 한다. 세부적인 내용은 다음과 같다.

1. 체계적 문헌고찰 및 임상연구 문헌에서 제시되는 효과크기 지표의 개념과 장단점을 분석한다.
2. 체계적 문헌고찰, 임상진료지침, 의료기술 평가 등 대표적인 근거중심의학 활동에서 이루어지고 있는 이득-위험 분석을 조사한다.
3. FDA, CHMP 등 신약의 승인과 관련된 기관에서 이루어지고 있는 이득-위험 분석의 현황에 대해서 분석한다.
4. 여러 가지 정량적인 이득-위험 방법론의 내용과 장단점에 대해서 분석한다.
5. 주요 질환 1개에 대한 표준적 치료적 의료기술의 효과 크기에 대한 NNT mapping을 통하여 임상적 효과성의 참고범위(reference value) 제시한다.

2. 연구방법

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

2.1 치료효과 크기 지표 분석

EBM교과서, Cochrane handbook, review article, clinical epidemiology, epidemiology textbook 및 관련 문헌을 검색하여 고찰하였다.

2.2 효과크기에 따른 의사결정 방법론

체계적 문헌고찰, 임상연구 및 방법론 연구를 pubmed를 이용하여 검색하였고, FDA, CHMP(Committee for Medicinal Products for Human Use) 등 신약의 승인과 관련된 기관을 검색하여 관련 방법론을 고찰하였다.

2.3 1개 질환 치료효과크기 mapping

2.3.1 질환 선정 기준

국제적으로 수용도가 높은 최신의학 교과서, 국내외 임상진료지침 등에 표준적으로 권고되고 있는 주요 질환의 주요 치료법에 대한 일차 연구들의 효과 크기에 대한 참고범위를 제시하고자 1개 질환을 다음과 같은 방법으로 선정하였다.

첫째, 질병의 자연사가 급·만성적인 경과과정이 있고, 둘째, 의료결과가 증상완화에서 사망까지를 포함한 질환으로 셋째, 외부효과가 없는 특정 organ에 관계된 질환, 마지막으로 국내외 임상진료지침에서 제시한 표준적인 치료법을 갖고 있는 질환을 선정하였다.

2.3.2 질환 후보

먼저 5개 후보질환에 대한 최신 업데이트된 guideline 검색하였다. 내부 연구진의 토의를 통하여 질환선정 기준에 적합한 급성 관상동맥 증후군(acute coronary syndrome; 이하 ACS)의 SIGN 가이드라인¹⁾을 mapping하기로 확정하였다.

SIGN의 가이드라인에 제시된 ACS의 치료법으로는 초기 질병 관리(initial management), ST 상승 급성 관상동맥 증후군에 대한 재관류(Reperfusion therapy for ST elevation acute coronary syndrome), 비침습적 검사와 위험분류(Risk stratification and non-invasive testing), 침습적 검사와 혈관성형술(Invasive investigation and revascularisation), 조기 약제 치료(Early pharmacological intervention), 저산소증과 심장 쇼크의 치료(Treatment of hypoxia and cardiogenic shock), 환자 지지와 정보 충족(Patient support and information needs)의 총 7개의 치료법에 대한 내용이었다.

표 1. 5개 후보질환에 대한 최신 업데이트된 guideline 검색

질환	Guideline 출판기관
Acute coronary syndrome	AHA, NCCHTA
Asthma	GINA
Osteoarthritis	ARA
Stroke	AHA, ACCP
Brand new type 2 DM	ADA, SIGN

2.3.3 치료효과 크기의 mapping 방법

2차원 그래프(y축: clinical importance, x축: NNT)로 작성하기로 결정하였으므로 SIGN의 가이드라인에서 RRR(Relative risk reduction), ARR(Absolute risk reduction)로 제시된 경우는 아래의 NNT 전환식으로 재계산하여 맵핑하였다.

7가지 중재법 내에서의 각 세부 치료법들 간의 NNT를 x축에 맵핑하여 그 분포를 조사하기 위하여 환자는 크게 Non-ST elevation ACS와 ST elevation ACS로 구분하였으며, 위약이나 보존적 치료법과 비교된 치료효과크기 값만을 추출하였다. 의료결과는 한

1) Scottish Intercollegiate Guidelines Network. Acute coronary syndromes. 2007.
(<http://www.sign.ac.uk/pdf/sign93.pdf>, 접속일자: 2010. 1. 11)

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

중재법에 여러 가지 의료결과에 대한 치료효과크기 값이 제시된 경우 모두 추출하였다.

y축의 임상적 중요성(clinical importance) 또는 효용(utility)이라고 할 수 있는 ACS의 치료결과에 대한 각 의료결과는 관련 논문들을 참고하여 연구진 토의 후에 결정하였다. 그 결과 사망 0, 출혈 및 주요한 심장에 대한 부작용(Major adverse cardiac event; 이하 MACE)의 경우 0.7 ~ 0.8, 심근경색증과 뇌졸중 발생의 경우 0.5로 점수를 부여하여 맵핑하였다(Samsa 1998; van't Hof 1999; Sinclair 2001).

표 2. NNT 전환식

Risk difference에서 계산

$$NNT = \frac{1}{\text{absolute value of risk difference}} = \frac{1}{|RD|}$$

Risk ratios의 메타분석의 결과들로부터 absolute risk reduction이나 NNT를 계산

ACR (an assumed control risk)

$$NNT = \left| \frac{1}{ACR \times (1 - RR)} \right|$$

Odds ratios의 메타분석의 결과로부터의 absolute risk reduction or NNT를 계산

$$NNT = \left| \frac{1}{ACR - \frac{OR \times ACR}{1 - ACR + OR \times ACR}} \right|$$

3. 연구결과

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

3.1 치료효과 크기 지표 분석

3.1.1 치료효과 크기 지표 개념

치료효과는 RCT에서 발견된 어떤 결과들의 임상적 유의성을 평가하기에 유용한 정보를 제공하고, RCT 설계 과정의 검정력 계산에 기초하며, 메타분석에서 RCT의 결과들을 대표해야한다. 즉, 통계적으로 유의한 결과의 임상적 유의성을 판단하기 위해서 치료효과가 필요하다.

잘 설계된 RCTs에서 treatment와 control group간의 통계적으로 유의한 차이 ($p=0.05$, $p=0.01$, $p=10^{-6}$)가 있을 때, 이러한 결과들은 자동적으로 임상적 유의성으로 간주될 수 있는가? 대답은 그렇지 않다. P value는 통계적 유의성 검정에는 유용하지만 임상적 유의성을 추론하기에 불충분하다.

3.1.1.1 이분형 자료의 분석 (Dichotomous data)

사례 연구, 질문:

대장 선종 절제술을 받은 환자에서 선암 재발 방지를 위한 목적으로 어떤 특정 약제를 권고할 수 있는가?

1) 상대적 효과 측정(Measure of relative effect)

Risk ratio (relative risk)

정의: 두 그룹에서의 event의 risk에 대한 비율(ratio)임

Risk = number of events of interest / total number of observations

Risk ratio = intervention risk / control risk

▪ 장점

- 환자와 보건의료전문가에게 친숙한 개념
- 이분형 의료결과에 대한 메타분석의 수행 시 권유됨
- 치료법에 대한 risk ratio

$$RRR(\text{Relative risk reduction}) = 100\% \times (RR - 1)$$

- absolute effect measure보다 더욱 일관적임

- odds ratio대신에 percentage reduction으로 표현할 수 있음
- 사건발생률이 희귀하고 치료법이 과히 성공적이지 않다면, RR은 OR에 근접함
- 중재군과 대조군에서 no events나 all events가 관찰된다면, 연구는 event의 상대적인 확률에 대한 정보를 제공하지 못함

■ 단점

- 대조군에서 사건이 발생하지 않는다면, 0으로 나눌 수 없기 때문에 계산될 수 없음
- 중재군과 대조군의 event proportion에 대한 absolute difference(magnitude)나 control group event rate에 대한 정보를 제공하지 않음
- 비대칭적인 특징을 가짐
- baseline event rate에 따라 달라지므로 다른 event rate을 가진 연구들의 메타 분석은 문제가 있음
- Montreuil(2005)이 제시한 예를 통해 보면 어떤 새로운 약이 수술 후의 부작용을 줄이는 지 무작위배정 비교임상시험을 하였다고 가정하고 약제 사용군과 대조군에서 수술 후 합병증 발생이 <표 3> 과 같았다면 새로운 약제의 효과를 측정하기 위해 수술 후 합병증의 위험의 감소효과는 'RR of Cx(+)'와 같이 0.5 값을 보여 대조군에 비해 수술 후 합병증 발병의 상대적 위험은 50%인 것을 알 수 있다. 하지만 수술 후 합병증이 발생하지 않을 상대 위험은 어떤가? 발생의 위험이 50%이면 발생하지 않을 위험은 두 배여야 할 것 같지만 실제 RR값을 계산하면 'RR of Cx(-)'으로 1.02 배가 된다. 즉 RR값은 어떤 사건이 일어날 확률이 높은 경우에는 작은 변화를 보여주고 상대적으로 적을 '그 사건이 일어나지 않을 경우'에서는 큰 값을 나타내게 된다. 치료 효과에 대한 크기의 인지는 성공률이나 실패율 중 연구자가 임의로 선택하여 보고할 때 영향을 받을 여지가 있게 된다. 이에 반해 Odds Ratio의 경우에는 표에서 보는 바와 같이 한쪽 결과 값의 OR의 역수가 상대측 결과 값의 OR의 역수가 되므로 이러한 현상은 사라지게 된다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

표 3. 새로운 가상의 신약의 수술 후 합병증 발병에 미치는 영향에 관한 가정 예

	Post-op Cx (+)	Post-op Cx (-)	Cx. risk
new Med	2	98	0.2
control	4	96	0.4
RR of Cx(+); $0.2/0.4 = 0.5$			
RR of Cx(-); $0.98/0.96 = 1.02$			
OR of Cx(+); $(2/98)/(4/96) = 0.49$			
OR of Cx(-); $(98/2)/(96/4) = 2.04$			
* post-op Cx(+); 수술후 합병증 발생함, post-op Cx(-); 수술후 합병증 발생하지 않음.			

Odds ratio

Odds ratio는 이분형 의료결과(성공/실패)에 대해, 가장 최근에 일반적으로 사용하고 있는 효과크기는 odds ratio(이하 OR)이다. OR은 어떤 정해진 threshold를 초과하는 모든 OR을 발견하기 위한 80% 이상의 검정력을 가지기 위해서 충분히 큰 표본수를 결코 제시해줄 수 없다. OR은 원래 우연성(randomness)에 대한 귀무가설을 검정하기 위해 도입되었고, 여전히 성공률이 우연이 아니게 다른지를 확인하는 민감한 지표로 남아있으며, 이는 질환의 유병률이 알려져 있지 않은 후향적 환자대조군 연구에서 널리 사용되고 있다²⁾.

정의: event의 odds에 대한 비율

odds = number of events of interest / number without the event

odds ratio = control odds / event odds

장점

- gambler에게 더욱 친숙한 개념
- 이분형 의료결과에 대한 메타분석의 수행을 권고함
- absolute effect measure보다 더욱 일관적임
- 0부터 무한대까지 값을 취할 수 있으므로 수학적 특성이 우월함
- 대칭적인 특성을 가짐
- 가정된 대조군의 위험의 맥락에서, risk ratio로 전환하여 해석을 용이하게 할 수 있음

$$\text{Risk ratio} = \frac{OR}{1 - ACR \times (1 - OR)}$$

ACR (assumed control risk = 연구들에서의 median control group risk)

2) Lipsey, Wilson. Practical Meta analysis.

(http://psych.unl.edu/psycrs/971/meta/effect_sizes.pdf, 접속일자 2010. 1. 11)

■ 단점

- 임상에서 적용하고 이해하기에 가장 어려운 요약 통계량임
- 중재군과 대조군에서 no events나 all events가 관찰된다면, 연구는 event의 상대적인 확률에 대한 정보를 제공하지 못함
- 상대위험도로 오인하는 경우가 있음. OR값을 보고한 5개의 외과학술지에 1999년에서 2004년까지 5년간 발간된 문헌 84개에 대한 분석 결과 15개 문헌(18%)에서 OR값을 위험도의 증가로 표현하는 오류를 범하였다(Montreuil 2005). 기저 위험률이 낮을 경우 OR은 RR 값에 근접하나 기저 위험율이 커지거나 작아지면 차이가 [그림 1]과 같이 벌어지게 된다.

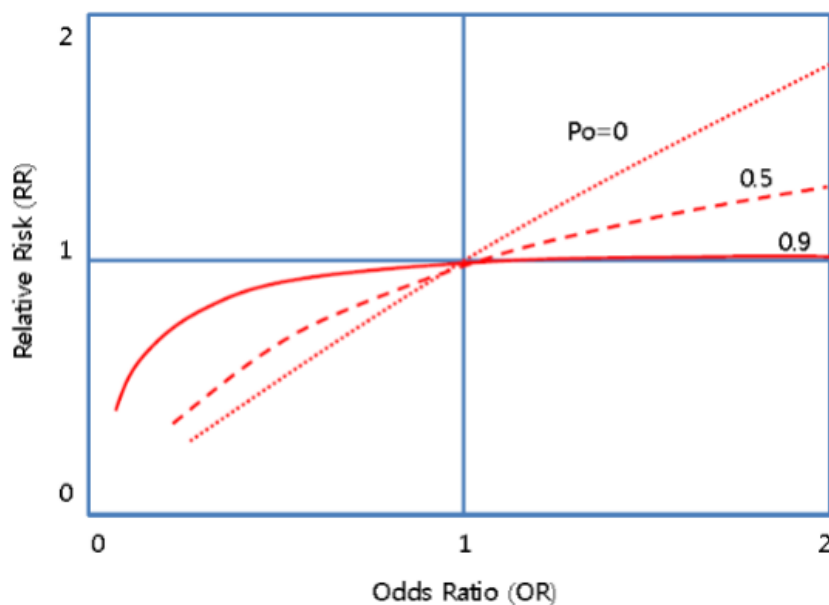


그림 1. 대조군에서의 결과발생의 빈도(P_o)에 따른 RR과 OR과의 관계
Modified from Montreuil 2005

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

사례연구: celecoxib의 선종예방 연구, Bertagnolli, 2006

상대적 특성을 지닌 지표

	3yr	
	adenoma	no adenoma
Celecoxib (배정된 수:679)	66	291
placebo (배정된 수: 685)	83	203
	RR=0.64(0.48-0.85)	
	RRR=0.363 (0.154-0.52)	
	ARR=0.105(0.039-0.172)	
	NNT=9.493(6-26)	
	OR=0.55 (0.38-0.80)	

대조군의 사건 발생 위험: X

치료군의 사건 발생 위험: Y

치료법의 사건발생의 상대위험도(RR): X/Y

치료법의 상대위험도감소(RRR): (1-Y)/X

(위약군에 비하여) Celecoxib는 선종 발생을 RRR=0.363 (0.154-0.52) 로 36% 감소시켰다.
Celecoxib는 선종 발생의 OR= 0.55 (0.38-0.80) 이다.

2) 절대적 효과 측정(Measure of absolute effect)

Risk difference

임상시험에서 두 그룹간에 평균값에서의 절대적인 차이를 측정한 표준화된 통계량으로 모든 연구들에서 의료결과의 측정법이 동일한 scale일 때, 메타분석에서 요약 통계량으로 사용될 수 있다.

장점

- 해석하기에 쉬움
- 어떤 한 집단에서 사건이 발생하지 않은 경우라도 계산될 수 있음
- 중재법의 likely benefits과 likely harms간의 trade-off를 고려할 때 유용함
- RD의 역수는 NNT이며, 이는 현장에서 임상에게 유용한 measure임
- 모든 trials의 결과들은 same units로 표현되고, same measurement tools로 획득되어 통합되어야 함

▪ 단점

- 임상적 중요성은 underlying risk of events에 따라 달라짐
- 결과를 기저 위험도가 다른 risk를 가진 환자집단에 적용시키기 어려움
- 중재효과의 일관된 추정치를 제공하지 않을 가능성이 있음
- 임상상황에서 risk difference가 일관적이지 않다면, 메타분석을 수행해서는 안됨

NNT(Moher 1998)

주어진 시간 frame에서 event를 피하거나 발생하기 위해 1명의 추가적인 사람 (additional person)에 대해 비교 중재법보다 실험 중재법을 받을 필요가 있는 사람의 기대되는 수로 정의된다.

$$NNT = \frac{1}{\text{absolute value of risk difference}} = \frac{1}{|RD|}$$

NNT는 risk difference로부터 도출되기 때문에, 그것은 여전히 효과의 comparative measure(experimental versus a certain control)로 single intervention의 일반적인 특성이 아니다.

OR 0.35나 effect size 0.42는 patient benefit의 용어에서 무엇을 의미하는가? 중재법의 cost, availability, adverse events에 대한 risk와 같은 particular outcome measure는 얼마나 중요한지, 새로운 치료법을 채택할 의사결정은 편익의 정도에 달려있다. dichotomous statistical measures를 좀 더 의미있는 measure로 전환시키는 한 가지 방법은 NNT concept이다.

NNT는 독자에게 관심있는 1개의 event를 만들기 위해 새로운 치료법으로 치료받을 필요가 있는 환자 수는 몇 명인지를 알려준다.

결과들이 summary risk difference로 $NNT = 1 / (\text{absolute risk difference})$ 이다.

결과들의 OR로 표현된다면, NNT의 계산은

$$NNT = \frac{1 - [PEER \times (1 - OR)]}{PEER \times (1 - OR) \times (1 - PEER)}$$

(PEER= patients' expected event rate)

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

NNT가 adverse events에 대해 사용될 수 있으며 이러한 경우 1 adverse event를 산출하기 위해 치료받을 필요가 있는 환자 수는 얼마인지를 알려준다.

개념 도입의 비교적 초기에 한 사건(事件; event)을 막기 위해 얼마의 노력을 들여야 할지보다 더 명확히 표현하기 때문에 임상 의사들로부터 이해하기 쉬워 매우 유용할 것으로 추천되었고(Laupacis 1988), CONSORT(Consolidated Standards of Reporting Trials)에서 무작위 대조군 임상시험의 결과를 보고할 때 NNT값을 보고하도록 권하며(Altman 2001) 근거중심의학 관련 이차간행학술지와 정책결정자들의 요구가 늘고 있다(McAlister 2008). 유럽의 규제당국은 NNT값을 각 임상연구에서 편익과 위험을 분석하기 위한 측정값으로 인지하고 있다(CHMP 2007).

adverse event에 대한 상대적인 benefit은 새로운 중재법의 채택여부를 결정하기 위해 평가될 수 있다.

threshold를 확립하는 1가지 접근법이 있을지라도, 누군가가 중재법을 명확하게 채택해야만 하는지를 의미하는 NNT의 절대적인 cutoff value는 없다. 개별 환자에서, 정보는 환자들이 이해할 수 있는 형식으로의 전환이 필요하며, 그래서 informed decision을 할 수 있어야 한다.

Treatment specificity

NNT는 치료법에 특이적이며, 특별한 임상적 의료결과를 도달하는 데에 대조군과 active treatment간의 차이를 기술한다.

NNT 1은 우호적인 의료결과가 치료가 주어진 every patient에서 발생하고 비교집단에서 no patient로 발생하는 것을 의미한다.

NNT 1은 치료가 주어진 모든 환자들인 특별한 의료결과를 성취할 것을 의미하며, 이것은 'ideal'한 것이다. 대부분의 예방적 치료전략은 환자집단 모두에서가 아니라 일부 환자집단에서 효과적이기 때문에 1보다 NNT가 크다.

올바르게 특이적인 NNT는 항상 비교자, 치료에 대한 의료결과, 의료결과를 도달하기 위해 필요한 치료기간, 95% CI를 제시해야만 한다.

Calculating NNTs³⁾

$$NNT = 1/(\text{중재군에서 편익이 되는 비율} - \text{대조군에서 편익이 되는 비율})$$

3) Bandalier. What is an NNT?, Implementing NNTs.

치료군에서 매우 적은 사건이 발생하는 예방법에서 계산은 음수 NNTs가 나오게 된다. 이때 부호는 무시하면 된다.

매우 효과적인 치료법의 NNT는 일반적으로 2-4 범위이다.

많은 수의 인구집단에서 영향을 받아 적은 수의 환자가 발생하는 예방법의 경우 더 큰 수의 NNT는 유용하다.

NNT 결과는 분리하여 사용되어서는 안 된다. 우리는 연령, 역학, 인구집단의 요구, 사회적 요소와 다른 지역적인 우선순위를 고려할 필요가 있다.

▪ NNT의 장점

- clinical and purchasing decision에 대한 정보를 주기 위해서 치료적 영역간과 내에서 효과성을 결정하기 위한 currency로써 사용될 수 있음
- NNT는 NNH로 계산될 수 있음: 새로운 중재법에서 안전성에 관한 이슈의 반영이며, NNH는 보통 adverse events라고 불리는 한 사람이 해(harm)가 되기 위해 치료받아야만 하는 사람들의 수를 반영함
- 그러므로, 임상적 위험 : 편익비의 평가는 $NNH : NNT \text{ ratio}$ 를 통해 기술될 수 있음
- $NNH : NNT \text{ ratio} < 1$ 이면, 임상적 위험 < 임상적 편익
- 상대적 특성을 갖는 측정법의 한계를 보완할 수 있음

어떤 치료법이 다양한 상태에서 이득이 어떻게 발생하는지 측정하는 것에 있어서 RRR과 같은 지표가 갖는 한계가 있다. 가상의 두 치료법을 예로 들어 비교해 보자(표 4). A 치료법의 경우 100명이 치료 받은 사람 중에 10명에서 관찰하고자 한 좋지 않은 결과가 발생하였고 대조군에서는 100명중 20명이 발생하였다고 가정해보자. B 치료법의 경우 10,000이 치료 받았고 이중 100명에서 좋지 않은 결과가 발생하였고 대조군에서는 10,000명중 200명에서 발생하였다고 하면 두 경우 모두 RRR값은 0.5로 A와 B치료 모두 위험을 50% 줄였다. 그러나 ARR값으로 보면 A치료는 대조군에 비해 위험발생률의 차가 10%나지만 B치료의 경우에는 1%밖에 나지 않으며 NNT 값으로 보면 A치료는 10으로써 10명을 치료할 때 한명에서 위험을 피할 수 있으며 B치료의 경우에는 NNT값이 100으로써 100명을 치료할 때 한명에서 치료의 이득을 보게 된다(Kumana 1999). RRR 값으로 판단할 경우 두 치료의 효과는 같게 여겨질 수 있지만 절대적 개념을 지닌 결과 측정법을 사용하면 두 치료의 효과는 확연히 차이가 난다. NNT 값을 사용하면 치료법의 비용과 치료법으로 인한 부작용의 정도와 빈도를 고려하여 득과 실을 판단하는 데 더 도움을 받을 수 있다. 즉 두 경우의 치료법에 대한 가치를 같게 볼 수는 없으

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

나 RRR 값으로 말하면 효과는 같아 보인다. NNT로 측정할 경우 효과를 보지 못하면서 치료받게 되는 환자의 규모, 그 비용, 또한 치료에 따른 부작용의 규모들을 고려하면 보다 합리적인 결정을 할 수 있는 기회를 갖게 되는 장점이 있다.

표 4. 가상의 두 치료법에 대한 여러가지 효과 측정방법에 따른 결과값

		사건(event)발생	사건 발생 하지 않음
A 치료법	A 치료군	10	90
	대조군	20	80
	RRR% (95% CI)	50(37.5-60)	
	ARR% (95% CI)	10(6.91-13.09)	
	NNT (95% CI)	10(7.634-14.491)	
B 치료법		사건(event)발생	사건 발생 하지 않음
	B 치료군	100	9,900
	대조군	200	9,800
	RRR% (95% CI)	50(36.5-60.6)	
	ARR% (95% CI)	1 (0.663-1.337)	
	NNT (95% CI)	100(74.8-150.7)	

- 동일 질환에서 NNT값을 통해 여러 치료들을 비교해 볼 수 있음

NNT는 치료-특이적으로 상대적 효능의 비교에 유용하다(Oxford 2007). Andrew Moore, Henry McQuay 등의 Oxford대학의 과학자들로 구성된 Bandolier는 근거중심보건의료에 관한 독립적 학술지 성격의 electronic version of Bandolier(www.ebandolier.com)를 운영하고 있고 진통제에 대한 대진표를 제시하고 있다(부록 1). 그중 일부를 도식화한 것은 [그림 2]와 같다.

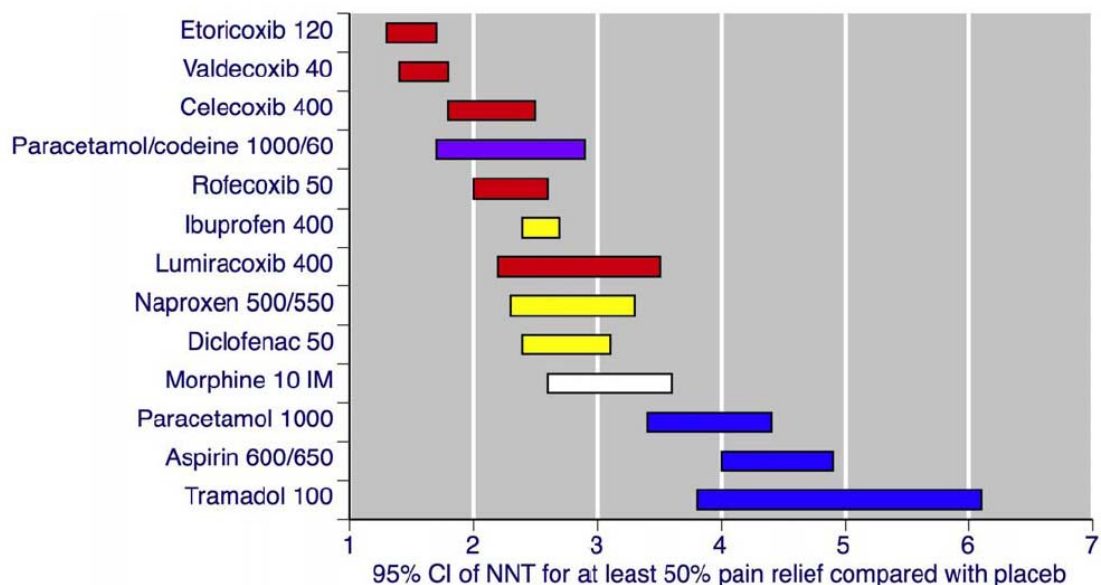


그림 2. 중등도 이상의 통증에서 복용 후 4-6시간 지나서 동통의 50%가 감소되는 것을 효과로 보았을 때 NNT값의 대진표

그림 3에서는 류마티드 질환분야에서 NNT값과 NNH값을 함께 제시하여 효과의 비교를 돕는 시도를 하였다. 그러나 이 문헌에서 대조군이 다른 경우 즉 위약군이 아닌 MTX나 SSZ의 경우를 위약군과 비교한 경우와 NNT값을 직접 비교하거나 결과 측정변수가 다른 즉, ACR과 mean change RJC, Paulus 와 같은 서로 다른 지표의 효용에 대한 보정을 하지 않고 직접 비교하는 것은 오류를 범할 수 있어 주의를 요한다.

골다공증에서도 약물 복용 3년 지난 경우 척추골절을 예방하는 무작위배정 비교임상시험이 이루어진 약제들 간의 비교에서 NNT값이 Strontium ranelate가 9, zoledronate가 14, alendronate가 15, ibadronate가 21의 결과 값을 보인 것을 보고하였는데 (Ringe 2009) 포함된 문헌에서 기저위험의 정도와 같은 변수에 대해 차이가 나는지 살펴보고 값의 차이에 대한 해석을 요한다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

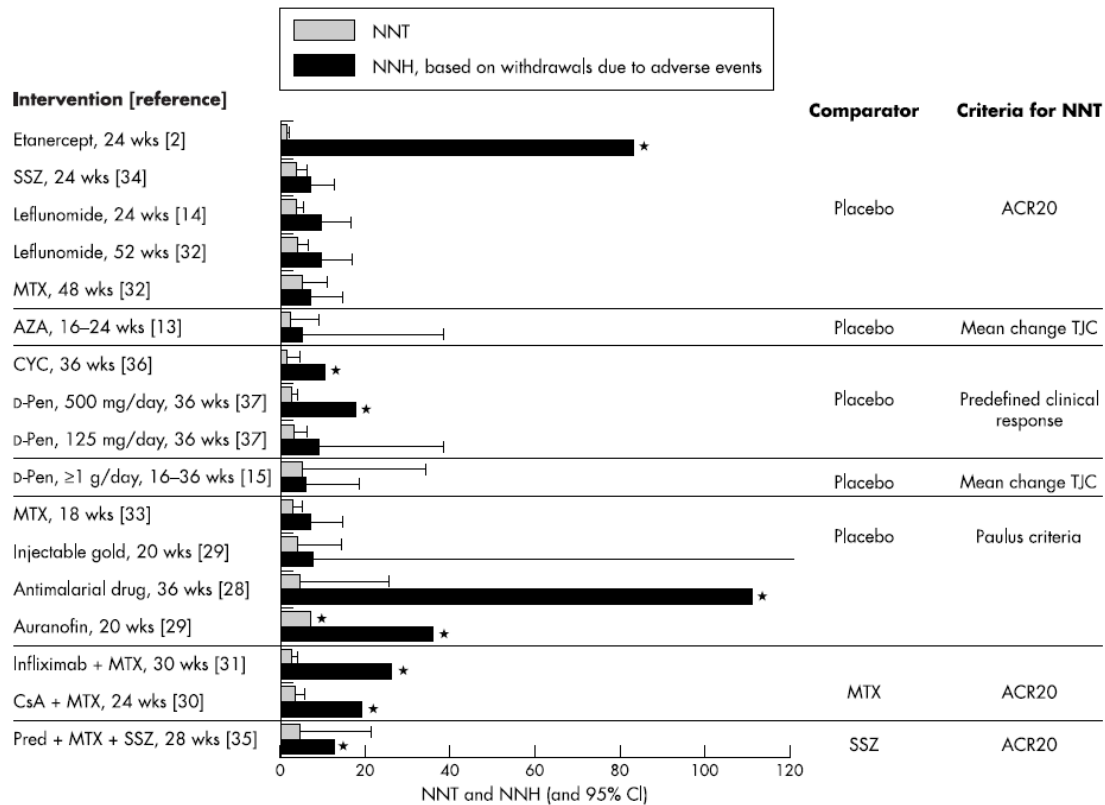


그림 3. 임상시험과 체계적문헌고찰에서 얻어진 NNT와 NNH

* 95% CI not calculated for non-significant results. SSZ, sulfasalazine; MTX, methotrexate; CYC, cyclophosphamide; AZA, azathioprine; D-Pen, D-penicillamine; CsA, cyclosporin A; Pred, prednisolone; TJC, tender joint count. (Osiri 2003)

■ NNT값에 영향을 끼치는 인자

- 기저 위험의 정도

위험의 정도가 다른 아군(subgroup)간에 어떤 치료법의 효과를 보는 상대적 지표(RR 혹은 RRR과 같은)가 큰 차이를 보이지 않는다고 할지라도 NNT는 기저 위험의 정도에 역상관관계를 갖게 된다. 낮은 위험도를 갖는 군에서는 NNT값은 좀처럼 이롭게 나오기 힘들게 된다(McAlister 2008). 이러한 점이 오히려 어떤 기저 위험도가 있는 환자에서 치료가 정당화 될 수 있는지를 고려하는 방법으로 NNT의 역치에 대한 개념을 도입하게 된다(Sinclair 2001). 절대위험도 감소나 이익의 역수인 NNT가 기저 위험 정도에 영향을 받는 것에 비해 Odds Ratio나 Relative risk는 기저 위험 정도에 상당히 독립적인 편이다(Cates 2002).

- 대조군의 종류

위약을 투여한 군과 기존의 가장 효과적인 치료를 투여한 대조군과 비교한 경우 후자의 경우 보고자 하는 위험이 발생할 위험이 줄어들므로 NNT값은 커지게 된다. 즉 대조군에 행하여지는 치료에 따라 대조군에서 발생하는 위험이 줄어들어 치료군에서 발생하는 위험률과 차이가 적게 나게 될수록 NNT값은 커지게 되므로 특히 연구 간의 비교나 치료법간의 비교를 할 때 대조군에서 행하여진 치료법이 무엇이며 대조군에서의 위험률이 어떤지를 주의 깊게 보아야 한다.

- 결과 측정에 소요된 시간

장기간에 걸친 결과측정이 이루어질 경우 상대위험도감소(relative risk reduction: RRR)에 큰 차이가 나지 않더라도 보고자하는 결과의 위험발생이 증가함에 따라 NNT 값이 감소할 수 있다. 지질강하제인 스타틴의 효과에 대한 연구 중 하나인 ASCOT-LLA 연구에서 90일의 NNT는 364 (95% CI 210~362)이었으나 연구의 3년 후 최종 결과 시점에서는 93 (95% CI 59~08)이었다(McAlister 2008).

동 문헌에서 McAlister는 NNT가 치료의 효과를 보고자 하는 위험의 발생을 완전히 피하는 것보다는 지연시키는 고혈압이나 골다공증 같은 만성 질환에서 보다 장기적인 후유증을 남기지 않는 상부 위장관 출혈과 같은 급성인 질환의 상태에 더 잘 적용이 될 수 있으며 전자와 같은 만성적 상태에 대한 예방적 치료의 효과를 보는데 있어서는 한 시점에서의 생존의 차이를 보는 것보다는 평균생존기간 같은 결과값이 더 낫다고 주장한다. 하지만 문외한 입장에서는 평균생존기간에서 이득의 증가보다는 NNT값이 더 이해하기 쉽고 치료를 더 잘 받아들이는 것으로 나타난 연구 결과도 있다(Halvorsen 2007).

■ NNT의 단점

- 시간(e.g. 1년 또는 6개월?)와 의료결과(e.g. 사망만 또는 모든 관상동맥 사건?)의 용어에서 표준화된 단위가 없음. 이것은 연구결과들은 다른 시간대와 다른 의료결과의 병합을 제공하고 있음을 의미함
- 여러 다른 의료결과들이 위치해있는 가치와 효과성의 고려에 대한 cut-off가 없음. 생명연장 치료는 더 높은 NNT가 수용가능한가? 이러한 딜레마에 관한 합의가 없어 현재시점의 해석은 그러므로 개별적인 편향에 여전히 노출되어 있음
- placebo는 치료에 대해 약간 긍정적인 경향을 포함할 수 있기 때문에, placebo-controlled trials로부터 계산된 NNTs는 과대추정 될 수 있음을 인지하여야 함

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

- NNT의 사용은 다른 정보와 요구로부터 별개로 사용될 수 있음. 지식은 반드시 임상 실무가 되기 위해서 경험과 관찰과 결합하여 사용되어야만 함
- NNTs는 효과적인 보건의료를 제공하는데 에서 단지 한가지의 고려사항임; 그것들은 다른 정보원과 함께 사용될 필요가 있음

■ 비교 시 제한점

여러 치료방법들 간에 비교를 하고자 하는 노력과 대진표(league table)를 만들어 동일 질환 상태에 대체 가능한 치료법들의 효과를 비교하기 위한 시도들이 있다(Osiri 2003 Gray 2005. Oxford 2007). 그러나 앞서 지적한 한계점들로 인해 유사한 비교군, 동일한 결과 지표, 위험의 정도가 유사한 환자군, 결과 값 측정시점이 유사한 연구에서 얻어진 결과 값이 아닌 경우에 NNT값을 직접 비교하기에는 무리가 따른다. 임상적으로는 납득하기 곤란하게 acetaminophen과 morphine과 같이 확실히 효과가 다른 이질적인 약제의 NNT 값이 크게 다르지 않게 관찰되는 것처럼 다른 환자 그룹 간에 비교한 결과값으로 직접 비교하는 것에는 문제가 있을 수 있다(Gray 2005).

■ 메타 분석에서의 주의

메타 분석에서 얻어진 NNT값은 연구 결과들 간에 절대위험도감소의 변이만이 아니라 무작위 대조군 시험의 대조군에서 기저 발생율에 마찬가지로 영향을 받게 된다는 점과 개인에게 적용하게 될 때 환자의 위험정도를 고려해야 한다는 점에 주의해야 한다(Marx 2003).

Marx는 관상동맥질환환자에서 n-3 polyunsaturated fatty acids의 효과에 대한 연구를 예를 들어 메타분석 결과 사망률에 대한 RRR은 19% (CI 10%~27%)로 효과적이었으며 합동(pooled) 절대위험도감소에서 유추한 NNT 값은 73(CI 49~147)이었으나 메타 분석에 포함된 연구들에서 기저 위험률이 2~22%였고 합동 상대위험(pooled RR)으로부터 유추된 NNT값은 기저 위험률이 2%인 경우 263(CI 185~500)이었고 22%였던 경우 24(CI 17~45)로써 기저 위험률에 변이가 큰 경우 합동 절대위험도감소에서 얻어진 NNT값은 정보력이 떨어지며 주어진 개인에게 적용할 때 오도될 수 있다는 점을 언급한다. 이러한 한계를 극복하기 위해 상대위험도감소의 메타분석 값과 적용하고자 하는 환자의 기저위험을 통해 NNT를 구하는 것이 더 적절한 방법이다(Smeeth 1999).

대조군에서의 위험정도를 알면(baseline risk BR) 다음과 같은 공식으로 RR값이나 OR값으로부터 NNT 값을 구할 수 있다(Wen 2005).

$$NNT = \frac{1}{BR(1 - RR)}$$

$$NNT = \left\{ \frac{1 - BR(1 - OR)}{(1 - BR)[BR(1 - OR)]} \right\}$$

또한 연구결과 밝혀진 BR 값과 실제 환자의 risk를 알면 환자의 요인 f를 구하고 이로 NNT값을 보정하는 방법이 제시되기도 하였다.

f = 환자가 갖고 있는 위험 / NNT값이 구해진 임상연구나 메타분석에서의 BR

$$\text{Adjusted NNT} = \frac{NNT}{f}$$

위에서 언급한 것과 같이 측정시기에 따라 NNT값이 변동 되므로, 해마다 해당 사건이 발생이 일정하다면 adjusted f를 사용하여 adjusted NNT값을 다시 구할 수 있다. adjusted f는 다음과 같이 구할 수 있다.

adjusted f = NNT값이 구해진 임상연구나 메타분석에서의 추적 기간 / 해당환자의 위험률을 구한 연구의 추적 기간

■ 연속형 변수에서의 사용

연속형 변수나 범주형 변수의 경우 일정 기준으로 이분형 변수로 전환을 하면 NNT값을 구할 수 있다. 최근 만성 요통등과 같이 환자가 보고하는 증상의 척도와 같은 범주형 변수에서 NNT값을 전환하는 시도가 있지만(Froud 2009) 아직 보편적인 방법은 아니며 일반적으로 NNT값은 사망, 심근경색 발생과 같은 이분형 변수에 국한되어 구할 수 있다는 제한점을 가지고 있다.

■ 신뢰구간

NNT값의 신뢰구간은 절대위험도감소(absolute risk reduction)의 신뢰구간의 역수를 취하여 제시할 수 있다. 만약 어떤 치료법이 대조군과 비교하여 절대위험도감소가 15%(95% CI 10~20%)라면 NNT값은 6.7(95% CI 5~10)이 된다. 그런데 절대위험도의 신뢰 구간 내 값이 0을 포함하는 경우 즉 치료의 효과가 전혀 없는 경우에는 그 역수는 무한대에 수렴하게 되므로 문제가 된다. 따라서 통계적 유의성이 없는 경우 혼

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

이 NNT값만 제시하고 신뢰구간은 제시하지 않게 된다. 이에 대해 Altman은 이런 문제를 극복하여 NNT의 신뢰구간을 제시하는 방법과 메타분석에서 도식화 하는 방법을 제시함으로써 어떤 경우에서라도 신뢰구간과 함께 NNT를 제시하는 방법을 제안하였다 (Altman 1998). 어떤 치료의 절대 위험도 감소가 0.1 (95% CI - 0.05 to 0.25)이라면 NNT값은 10(95% CI -20~4)가 아니라 95% CI의 값은 4에서 ∞ , 그리고 ∞ 에서 -20으로 표시할 수 있으나 NNT에 음수는 개념상 존재할 수 없으므로 이를 몇 명을 치료할 때 한명에서 해로움이 발생하는가의 개념인 NNTH로 표시하고 NNT도 몇 명을 치료할 때 한명의 이득이 발생하는가의 개념으로써 NNTB로 표시하게 되면 NNTB 10(NNTH 20~ ∞ , NNTB 4~ ∞)으로 표시할 수 있게 되고 이를 NNTB 10(NNTH 20~ ∞ ~NNTB 4)으로 표시할 수 있다. 또한 어떤 치료의 절대 위험도 감소가 -0.08(95% CI - 0.27~0.11)이라면 NNT값은 12.5(95% CI -3.7~9.1) 이 NNTH = 12.5(NNTH 3.7~ ∞ ~NNTB 9.1)으로 표시할 수 있게 된다. 코크란 그룹에서도 NNTB와 NNTH의 이러한 개념을 받아들이고 있다(Higgins 2008).

▪ 결과 값의 의미; 효용

다른 결과 측정값과 마찬가지로 수자 자체가 효용(utility)을 내포하지 않고 있으므로 사망률을 본 결과와 장애 발생률을 본 결과와 중세의 완화 정도를 본 결과에 따른 NNT값으로 동일선상에서 비교할 수 없다. 또한 해당 치료법의 비용 이에 따른 부작용의 발생 정도와 그 발생 부작용의 위중도 또한 고려해야 하므로 NNT값만 가지고 치료법을 비교하는데에는 무리가 따르게 된다.

사례연구: 절대적 특성을 지닌 측정방법

대조군의 사건 발생 위험: X

치료군의 사건 발생 위험: Y

치료법의 절대위험도 감소(ARR): $X - Y$ 필요치료수 (NNT)= $1/(X - Y)$

(위약군보다) Celecoxib의 투여군에서 선종 발생율이

ARR= 0.105(0.039-0.172)로 10%차이가 났다.

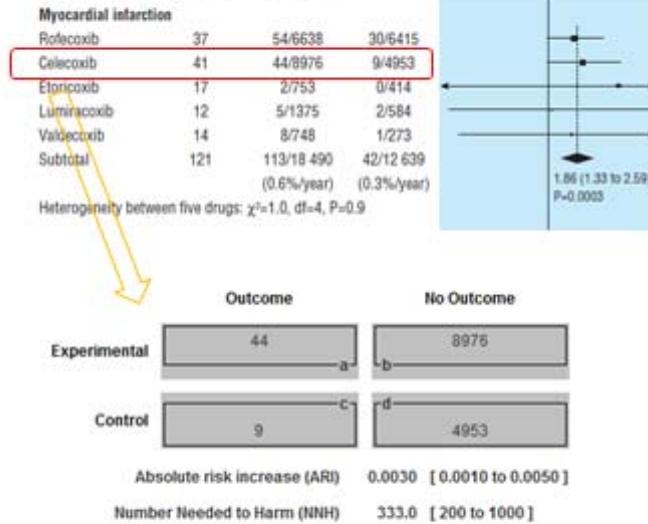
NNT=9.493(6-26)로써 3년간 9명이 Celecoxib를 투약받을 때 (3년 째에)
한사람에서 선종 발생을 예방하는 효과가 있게 된다.

NNT



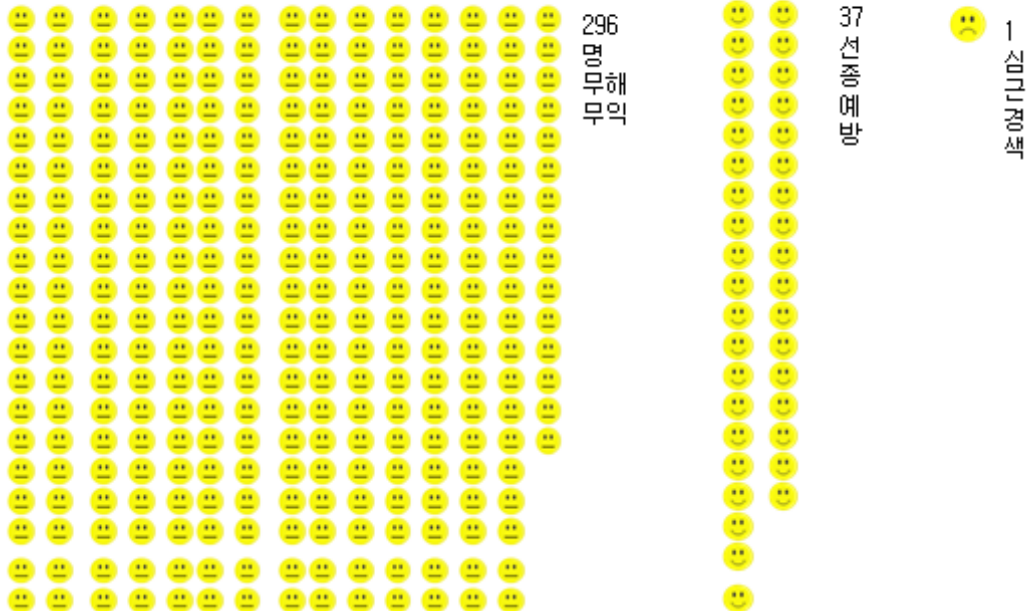
치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Do selective cyclo-oxygenase-2 inhibitors and traditional non-steroidal anti-inflammatory drugs increase the risk of atherothrombosis? Meta-analysis of randomised trials



Kearney, BMJ, 2006)

NNT=9.493(6-26), NNH=333(200-1000)으로써 3년간 9명이 Celecoxib를 투약받을 때 (3년 내에) 한사람에서 선종 발생을 예방하는효과가 있게 되며 333명을 치료할 때 이로인해 1명의 급성심근경색증이 생기게 된다. 333명 치료할 때 37명에서 선종이 예방되고 1명이 약제로 인해 심근경색



3.1.1.2 연속형 의료결과의 효과 측정 (Effect measures for continuous outcomes)

1) Cohen's d

가장 일반적으로 사용되는 효과크기인 Cohen's d와 Odds ratio가 있다. RCT outcome measure가 scale일 때, 가장 일반적인 effect size는 Cohen's d로 대조군과 실험군의 그룹간 평균의 차이는 그룹내 표준편차로 나뉘며, 두 그룹간의 response가 동일한 표준편차를 가진 정규분포인 경우를 위해 설계된 것이다.

$d \text{ or } r = d / (d^2 + 4)^{1/2}$, $2\Phi(d/\sqrt{2}) - 1$; Φ = 표준정규분포의 누적분포함수

$d=0.2 \rightarrow$ small effect size

$d=0.5 \rightarrow$ medium effect size

$d=0.8 \rightarrow$ large effect size

장점은 정규분포를 다루는데 익숙한 통계학자에게 Cohen's d는 표준정규분포 커브를 overlapping하는 이미지를 만들어주기 때문에 유익하지만, 단점은 d에 의해 결정되는 overlap의 정도가 임상가와 소비자에게는 임상적으로 해석가능한 메시지를 전달해주지 않는다는 점이다(Kraemer 2006).

Cohen's "Rules-of-Thumb"법칙에 의하여 SMD(standardized mean difference) 이외에도 상관계수(correlation coefficient)의 치료효과 크기의 해석은 다음과 같이 할 수 있으며,

0.10 $\rightarrow$ small effect size

0.25 $\rightarrow$ medium effect size

0.40 $\rightarrow$ large effect size

odds-ratio 의 해석은

1.50 $\rightarrow$ small effect size

2.50 $\rightarrow$ medium effect size

4.30 $\rightarrow$ large effect size

이라고 할 수 있다. 하지만 이것들은 중재법이 수행되는 맥락을 고려하지 않은 상태이기 때문에 해석에 주의를 기울여야 한다²⁾.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

2) Standardized risk difference

두 집단 간에 관찰된 risks(proportions of individuals with the outcome of interest)간의 차이를 표준편차(Standard deviation)로 나눈 값

■ 장점

- 대조군에 비해 중재군의 의료결과에 대한 평균 (또는 중앙값) 변화량 추정
- 중재법에 관한 효과성의 scale-free estimate를 제공하며, 다른 outcome measures를 사용한 연구들의 통합을 가능하게 해줌
- MD로부터 구해진 결과들은 일차연구들에서 사용된 measurement와 같은 단위에서 표현될 때, 해석하기에 쉬움

■ 단점

- scale의 방향에서의 차이를 수정해주지 않으므로 보정되어야 함
- 결과들을 통합하기 전에 단일 scale로 결과들을 표준화해야 함
- Z scores에서 표현된 결과들을 해석하는 것은 직관적이지 않음
- 모든 연구 결과들은 same units로 표현되고, same measurement tools로 획득되어 통합되어야 함(Mean difference)

3.1.1.3 Time-to-event data

Event가 경험되기 전의 시간경과에 초점을 두며, 통계학적으로 survival data로 알려져 있다. Time-to-event data는 각 개인에 대해 다음의 짝지은 관찰값으로 이루어진다((i) event가 관찰되지 않은 동안의 시간의 기간 (ii) event에 대응하는 time period의 끝이나 관찰의 끝의 indicator).

Event에서 끝나지 않은 어떤 시간의 기간을 기여한 참여자는 'censored'라고 불리며, 그들의 event-free time은 정보에 기여하여 분석에 포함시킨다. Time-to-event data는 때때로 이분형 자료(risk ratios, odds ratios, risk differences)로써 분석될 수 있으며, fixed time-point에서 알려진 한 연구에서의 모든 환자의 상태(status)가 요구된다.

time-to-event data를 요약하는 가장 적절한 방법은 생존분석 방법을 이용하고 중재 효과를 hazard ratio로 표현하는 것이다. 가정은 추적조사 기간에 걸쳐 hazard ratio는 일정하다는 것으로 이것은 비례위험(the proportional hazards)의 가정으로 알려

져 있다.

Hazard는 risk와 의미가 비슷하지만, 그것은 동시에 일어나는 위험을 측정하고 연속적으로 변할 수 있다는 점에서 다르며, 해석은 risk ratio와 비슷한 방식으로 hazard ratio는 그들이 control intervention보다 오히려 중재군의 중재법을 받는다면, 시간에 있어 특별한 시점에 연구대상자는 얼마나 많이 사건을 경험하는 지로 해석한다.

3.1.1.4 메타분석을 위한 효과크기 측정(Julian 2008)

메타분석에서 하나의 통계량을 사용하고 해석하기에 더욱 쉬운 통계량으로 다시 결과를 재표현하는 것은 합리적이다. 예를 들어, 메타분석은 relative effect measures (risk ratios, odds ratio)를 이용하고 그 결과들이 absolute effect measures (risk differences or numbers needed to treat)로 재표현하여 가장 잘 수행될 수 있다.

그러나 이들에 대한 모든 전환에서 experimental intervention이 적용될 것에서 'control' population에서의 outcome의 likely risk를 가리키는 baseline risk의 값을 상세화하는 것이 요구된다.

가정된 대조군의 위험도에 대해 선택된 값은 연구들에 걸쳐서 전형적으로 관찰된 대조군 위험도(observed control group risks)에 가깝고, absolute effect의 유사한 추정치는 메타분석에서 사용된 odds ratios나 risk ratios에 관계없이 구해질 것이다.

assumed control risk이 typical observed control group risk와 다르다면, absolute benefit의 예측은 어떤 summary statistic이 메타분석에 사용되었는지에 따라 달라질 것이다.

중재효과는 relative 또는 absolute effect로써 표현될 수 있으며, risk ratio (relative risk)와 odds ratio는 상대적 측정인 반면, risk difference와 number needed to treat는 절대적 측정이다.

메타분석을 위해 사용되는 summary statistic의 선택은 균형있는 3개의 기준에 의한다.

첫째, 우리는 메타분석에서의 모든 연구들과 중재가 수행될 인구집단의 하부그룹에 유사한 value를 제공하는 summary static을 원한다. 요약통계량이 더욱 일관될수록, single summary number로써 intervention effect를 표현하는 것은 더욱 정당화된다.

둘째, 요약통계량은 타당한 메타분석을 수행하기 위해 요구되는 수학적 특성을 가져야만 한다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

셋째, 요약통계량은 고찰을 이용하는 사람들에 의해 쉽게 이해되고 적용(의사소통)되어야 한다.

- 일관성: 경험적인 근거는 relative effect measures가 평균적으로 absolute measures보다 더욱 일관적임을 제시하였다. 이러한 이유로 특별한 임상상황에서 risk difference가 일관적일 것이라고 추정되는 명백한 이유가 없다면, risk difference의 메타분석의 수행을 피하는 것이 현명하다. 평균적으로, odds ratio와 risk ratio간의 일관성에서의 차이는 거의 없다.
- 수학적 특성: 가장 중요한 수학적 기준은 신뢰성 있는 variance estimate가 이용가능한지이다. number needed to treat는 다른 요약통계량으로부터 계산될 수 있을지라도, 메타분석에서 직접적으로 쉽게 사용될 수 없고 간단한 variance estimator를 가지지 않는다.
- 해석의 용이성: odds ratio는 임상에서 적용하고 이해하기에 가장 어려운 요약통계량으로 저자들은 메타분석에서 그것을 마치 risk ratios처럼 잘못 해석하여 출판한 경우도 있다. Systematic reviews의 결과들이 odds ratios로 일반적으로 표현되는 것은 결과들이 임상실무에서 적용될 때, 종종 치료법의 편익과 위해에 대해 과대추정을 유도할 것이라는 염려가 있다. Absolute measures of effect는 또한 relative effects보다 임상익에 의해 더욱 쉽게 해석될 수 있고, 중재법의 likely benefits과 likely harms간의 trade-offs를 허용할 수 있다.

표 5. Intervention의 효과를 표현하는 다음 5개의 효과크기: odds ratio, relative risk, risk difference, weighted mean difference, standardized mean difference¹⁾

효과 크기	정의	사용과 해석	제한점	장점									
		Discrete Data											
Odds ratio(OR)	Control group의 odds에 의해 intervention group에서 발생할 사건의 odds를 나눔 <table><tr><td></td><td>Dz(+)</td><td>Dz(-)</td></tr><tr><td>intrevnetion(+)</td><td>a</td><td>b</td></tr><tr><td>intervention(-)</td><td>c</td><td>d</td></tr></table> odds ratio = oddsdis / oddsnondis = 중재군에서 질병이 발생할 odds / 대조군에서 질병이 발생할 odds · 환자군에서 질병이 발생할 odds = a/(a+b) = a/b b/(a+b) · 대조군에서 질병이 발생할 odds = c/(c+d) = c/d d/(c+d) 따라서 교차비 추정치 OR = a/b c/d = ad / bc		Dz(+)	Dz(-)	intrevnetion(+)	a	b	intervention(-)	c	d	전통적으로 원하지 않는 사건(예시: 사망)을 줄이는 중재의 favorable effect는 OR < 1.0으로 표현됨. 그런 상황에서, OR > 1은 대조군이 favorable outocmes (예시: 불임 부부에서 임신)를 증진시키기 위해 주어진 중재법의 편익은 OR>1.0로 표현됨	clinical setting에서 쉽게 해석되지 않으며, 보통 잘못 해석되고 relative risk로써 간주됨. meta-analysis의 통계적 결과를 표현하는 data를 요약하기 위해 OR의 적절성을 염려하는 debate가 점차 있음	비교적 계산하기 쉽고, 지금까지는 의학논문에서 popular함
	Dz(+)	Dz(-)											
intrevnetion(+)	a	b											
intervention(-)	c	d											
Relative risk(RR)	intervention group과 control group 간에 사건이 발생할 상대적 확률을 표현. risk ratio와 event rate ratio로 알려짐	사건발생률이 낮을 때, RR과 OR은 유사하게 해석될 수 있음. clinical trail과 meta-analyses의 상황에서, RR의 complement(1-RR)은 relative risk reduction임	중재군과 대조군에서 사건의 proportion에서 absolute difference(magnitude)나 control group event rate에 대한 정보를 제공하지 않음	유사한 사건을 가지는 대조군에서 그들 환자들에 비해 사건을 가지는 중재군에서의 환자들의 상대적인 proportion에 대한 정보를 임상 의에게 제공. 임상 의에게 popular함									

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

효과 크기	정의	사용과 해석	제한점	장점
		Discrete Data		
Risk difference(RD)	대조군과 중재군에서 event rate간의 absolute difference로 정의됨	효과의 절대적인 측정값으로 중재군과 대조군에서 events의 frequency간에 absolute difference를 임상자에게 제공	임상의들은 효과의 absolute estimate보다 오히려 relative estimate를 통합시킨 효과의 측정치를 일반적으로 선호함	계산하고 해석하기에 비교적 쉬움. RD의 역수는 NNT이며 이는 현장에서 임상자에게 유용한 measure임
		Continuous Data		
Weighted mean difference(WMD)	중재군과 대조군간에 평균값에서의 weighted difference로써 중재효과를 표현	평균적으로 중재를 받은 환자들이 대조군에 비해 어떤 outcomes를 더 가짐(또는 덜 가짐)	모든 trials의 결과들은 same units로 표현되고, same measurement tools로 획득되어 통합되어야 함	WMD로부터 획득된 결과들은 그 것들이 일차연구들에서 사용된 measurement의 same units에서 표현될 때, 해석하기에 쉬움
Standardized mean difference (also called effect size)(SMD)	대조군의 SD(또는 pooled SD)에 의해 나뉘어진 중재군과 대조군간의 평균값의 absolute difference로서 중재효과를 표현	same outcome을 평가하는 trials들은 종종 다른 measurement tools을 사용함. 이들 상황에서 data를 사용하기 위해서 그들은 표준화된 단위로 전환되어야만 함 (SMD). WMD를 해석하기 위한 가이드라인들이 있음. SMD의 several variations는 SD의 선택에 달려있음	Z scores에서 표현된 결과들을 해석하는 것은 직관적이지 않음	계산은 비교적 쉽고, 그 결과는 SD units(z scales)로 표현됨 이것은 중재법에 관한 효과성의 scale-free estimate를 제공하며, 다른 outcome measures를 사용한 연구들의 통합을 가능하게 해줌

1) Moher et al. Guides for reading and interpreting systematic reviews. Arch Pediatr Adolesc Med. 1998;152(9):915-20.

3.1.2 치료효과크기 지표에 따라 의사결정이 달라진 사례연구

3.1.2.1 치료의 결과를 보고하는 방법에 따라 치료 효과의 정도를 인식하는 것에 영향을 주는가?

Naylor등은 Helsinki Heart Study (Frick 1987)에서 치명적 심근경색증과 모든 심근경색증 및 전체 사망률 세 가지의 최종 결과 값들을 상대위험도감소(RRR)와 절대 위험도 감소(ARR) 및 NNT 값으로 표현(예; '한명의 심근경색증을 막기 위해 5년간 77명이 치료받아야 함' 식의 표현)하여 의사 및 의과대학생을 대상으로 치료의 효과에 대한 판단을 내리도록 연구하였다(Naylor 1992). 중간값을 0점으로 하여 좌우 5점까지 표시하였고 좌우최대점은 해로움과 매우 효과적임으로 표시하여 제시된 결과값에 대한 판단을 하도록 하였다. 그 결과 RRR값으로 결과를 제시한 경우가 ARR로 제시한 경우 보다 효과 및 해로움 면에서 인식이 더 크게 나타났으며(그림 4), NNT로 표현한 경우보다 RRR값으로 표현할 경우 2.27점 높게 그리고 ARR로 표현한 경우가 1.80점 높게 인지되었다. 이 결과를 통해 볼 때 어떤 방식으로 연구 결과값을 제시하는가에 따라 효과에 대한 판단에 영향을 주었고 RRR 값으로 제시될 경우가 ARR이나 NNT로 제시하는 것보다 효과나 해로움에 대한 판단이 더 크게 인식되는 것으로 나타났다.

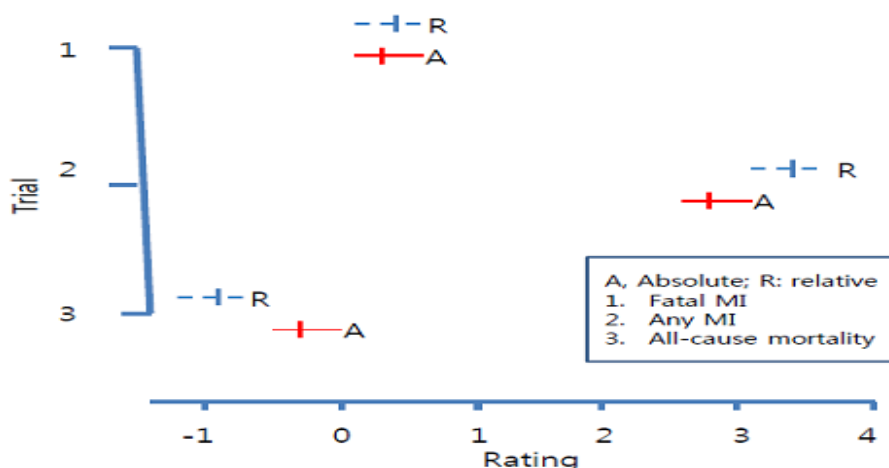


그림 4. ARR 및 RRR값으로 표현한 경우 효과에 대한 인식 정도
(Naylor 1992)

보건의료정책의 의사결정자들에 대한 연구에서도(Fahey 1995) 유방촬영술과 심장재활 치료프로그램에 대하여 상대위험도감소(RRR), 절대위험도감소(ARR), 무사건환자의 비율

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

(proportion of event free patients) 및 NNT값으로 제시하여 각 사람이 0(급여에 넣기 원치 않음)에서 10(급여에 포함하기를 매우 강하게 원함)까지 된 척도를 사용하여 자신의 판단을 표시하도록 한 결과 RRR값으로 제시한 경우에 제일 높은 점수를 보였다(표 6). 즉 어떤 측정값을 제시하느냐에 따라 의사결정자의 판단에 영향을 줄 수 있다는 것을 시사하는 연구 결과이었다.

표 6. 두 의료기술에 대한 급여의사결정에서 임상연구 결과 제시방법에 따른 정책결정자들의 판단. 평균 점수(%) (95% CI) (Fahey 1995)

데이터 제시 방식	유방촬영술	심장재활치료
상대위험도감소 relative risk reduction (RRR)	79 (76 to 83); (RRR=34%)	76 (72 to 80); (RRR=20%)
절대위험도감소 (ARR)	38 (35 to 42) $\neq$ T; (ARR=0.06%)	56 (53 to 60) $\neq$; (ARR=3%)
무사건 환자의 비율 (EFP)	38 (34 to 42) $\neq$ T; (EFP=99-82% v 99-8%)	53 (49 to 57) $\neq$ T; (EFP=84% v 87%)
NNT	51 (47 to 55) $\neq$; (NNT= 1592)	62 (58 to 66) $\neq$; (NNT=31)

* 더 높은 점수는 급여하기를 더 지지한다는 표현임.

Bonferroni difference means: $\neq$ significant difference at $P \leq 0.05$ from relative risk reduction;

T $\neq$ significant difference at $P \leq 0.05$ from number needed to treat

일반인을 대상으로 관상동맥질환의 위험을 감소시키기 위한 스타틴 약물을 복용할 지에 대한 의사결정을 약제의 효과를 상대적 지표인 상대위험도감소 그리고 절대적 지표인 절대위험도감소, NNT, 사건발생률, tablets needed to take 및 natural frequencies 여섯 가지 형태로 나타내었을 때 어떤 영향을 받는지 2,987명을 대상으로 한 무작위대조군연구를 수행한 결과 (Carling 2009) 상대적 척도인 RRR값으로 표현할 경우가 절대적 척도인 다른 지표들 보다 21% 정도 스타틴을 채택할 확률이 높았다(표 7).

표 7. 여러 다른 지표들로 제시된 사람들의 스타틴 복용 의사결정 (Carling 2009)

결과표시	RRR	ARR	NNT	ER	TNT	NF	Overall	p-value
스타틴 복용 시작결정, n (%)	376 (74)	256 (51)	267 (55)	252 (53)	284 (55)	256 (52)	1,691 (57)	<0.001

의료인, 정책결정자 및 일반인 모두 상대적 특성을 갖고 있는 RRR값이 절대적 특성을 갖는 ARR이나 NNT값에 비해 어떤 치료법에 대한 판단을 더 크게 인식할 수 있다는 것을 보여주고 있는데 이와 관련된 여러 편의 문헌을 분석하거나 체계적인 고찰을 한 후속 문헌들에서도 이와 같은 사실을 확인해 주고 있다(Moxey 2003, Covey 2007).

Covey는 치료의 효과의 측정법을 상대위험도감소와 절대위험도감소 혹은 NNT로 제시된 경우를 비교한 연구들로서 치료법의 효과에 대한 판단을 정량적 방법이나 혹은 선택 혹은 비선택과 같은 방식으로 살펴 본 연구들 중 각각의 측정법을 무작위 대조군시험으로 배정하여 수행하거나 같은 대상에게 여러 방법을 다른 시점에서 제시하여 측정한 연구들을 대상으로 한 체계적 문헌 고찰연구의 결과를 보고하였다(Covey 2007). 이 연구에서 31개의 개별적 연구들이 포함되어 분석 되었고 이질성을 보였지만 일관되게 상대위험도감소로 제시된 경우가 절대위험도 감소나 NNT로 제시된 경우보다 치료에 대해 더 옹호적인 평가를 받게 된다는 것으로 나타났다(그림 5).

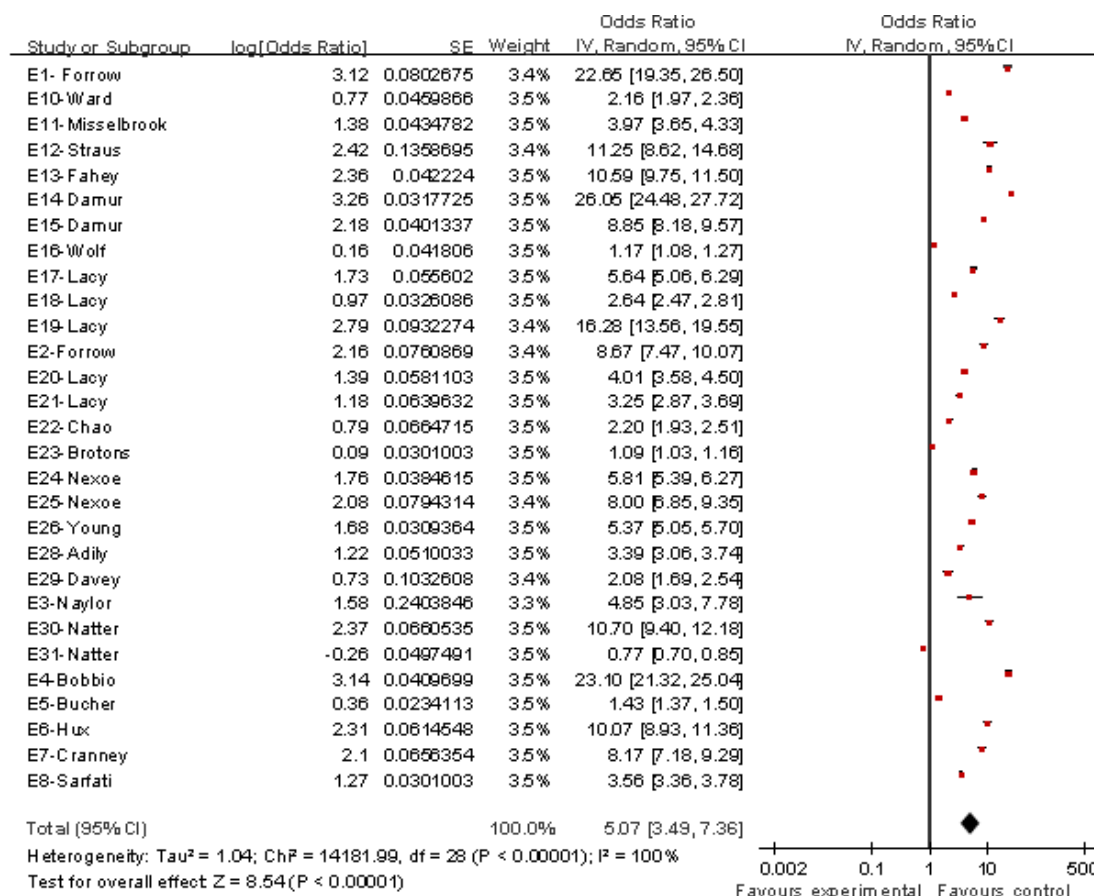


그림 5. 절대위험도감소 대비 상대위험도감소의 효과에 대한 판단의 비교

(Covey 2007 문헌에서 Generic inverse variance 데이터로 log odds ratio값과 confidence interval을 이용하여 계산한 SE값을 사용하여 random effect model로 메타 분석함. Favours Control이 RRR측임)

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

특히 많은 경우 이득에 관한 값은 상대적 성격의 결과 측정값으로 제시하고 부작용에 대해서는 절대적 성격의 결과 값들을 제시하는 경우가 있는데 이는 해로움 보다 이득이 더 크게 느끼게 할 수 있으므로 치우침을 발생시키게 된다(Wen 2005).

Covey의 연구 이후에 나온 문헌들에서도(Carling 2009, Griffith 2009) 이러한 현상은 일관성 있게 관찰되어 치료적 중재에 대한 효과의 측정 지표에 따라 효과의 크기를 판단하는데 있어서 어떤 측정법을 쓰느냐가 영향을 주므로 근거중심 의사결정과정에서 결과를 해석하고 가치판단(appraisal)과정에 이에 대한 고려를 해야 한다.

이상의 연구들을 통하여 본 결과 상대적 특성을 지닌 효과 측정방법인 상대위험도감소(RRR)는 절대적 특성을 지닌 효과 측정방법인 절대위험도감소나 NNT값에 비하여 효과에 대하여 더 크게 판정하게 되므로 RRR값만 단독으로 제시하는 것은 위험성이 있으며 RRR값과 더불어 절대위험도 감소(ARR) 및 NNT값을 제시하도록 해야 한다(Schechtman 2002).

3.2 효과크기에 따른 의사결정 방법론

3.2.1 여러 근거 중심 의사 결정에서 BRA

3.2.1.1 근거중심의학

Sackett 등은 근거 중심 의학은 5단계로 이루어져 있는데 그것은 1) 임상진료에서 생기는 의문점들을 대답할 수 있는 질문으로 형태화 하는 것, 2) 그 질문에 답이 될 수 있는 문헌을 찾는 것, 3) 그 증거의 타당도와 유용성을 세심하게 평가하는 것, 4) 그 결과를 임상에 적용하는 것, 5) 진료행위를 평가하는 것이라고 하였다(Sackett 1997).

근거중심의학에서 이득과 위험을 저울질하는 시기는 4단계로 연구 결과를 임상에 적용하기 위해서는 해당 중재의 이득과 위험에 대한 저울질을 하여야 한다. 제 4단계는 두 가지를 하여야 한다. 첫째는 해당 치료로 결과로 부작용보다 이득이 많은 가능성을 결정하여야 한다. 대부분 NNT(한 예의 부정적 사건을 예방하기 위해 치료해야 할 환자 수)와 NNH(한 예의 부정적 사건이 발생할 때 까지 치료를 받는 환자의 수)를 비교한다. 둘째는 환자의 선호도와 결합하는 것이다(Guyatt 1994).

3.2.1.2 체계적 문헌고찰

체계적 문헌 고찰은 1) 연구 기획단계, 2) protocol 작성, 3) 문헌 찾기, 4) 문헌 선정(포함-배제 결정), 5) 질 평가, 6) 데이터추출, 7) 데이터합성, 8) 보고서 쓰기(근거 해석, Trade-off, 결론 도출), 의 과정을 거쳐서 이루어진다⁴⁾.

체계적 문헌 고찰의 고찰에는 포함되어야 할 몇 가지 사항이 있다. 즉 해당 중재 효과와 안전성에 대한 근거의 강도, 리뷰의 잠재적인 비뚤림과 제한점, 결과의 적용성(일반화 가능성), 이득, 유해, 비용 사이의 저울질, 리뷰가 환자 진료 혹은 공중보건과 추후 연구에 미치는 영향(결론)에 대한 것을 기술한다. 이중 환자 진료의 공중보건에 미치는 영향을 결국 리뷰의 결론이라고 할 수 있으며 1) 고찰 결과 전혀 근거가 없다(Review found no evidence at all), 2) 고찰 결과 중재를 지지하는 명백한 근거가 있다(Review found evidence that clearly supports intervention), 3) 고찰 결과 이득이 없다는 명백한 근거가 있다(Review found clear evidence of lack of benefit), 4) 리뷰 결과 잠재적인 유해가 있다는 명백한 근거가 있다(Review found clear evidence of potential for harm), 5) 리뷰 결과 알려진 이득과 이상 반응 사이에 저울질이 필요하다는 근거가 있다 중 하나로 결론을 내리게 된다⁵⁾.

체계적 고찰에서 실제로 위험과 유해를 어떻게 저울질 하는지에 대해서 명백한 방법론이 알려져 있지는 않으며 결국 해당 질환에 대한 임상 지식과 근거중심의학의 원칙에 입각하여 정성적인 판단을 하는 것으로 알려져 있다.

다음의 경우를 생각해 보자.

일부 연구에서 아스피린 혹은 NSAIDs가 대장암 발생 위험이 중등도 정도인 경우 대장암 발생 위험을 낮춘다는 결과가 나왔지만 그렇지 않다는 보고도 있어서 이 주제에 대한 체계적 고찰이 필요하다고 판단하여 체계적 문헌 고찰을 시행하였다. 아스피린 혹은 NSAIDs의 대장암 예방암 예방 투여의 이득과 유해를 저울질 하였다.

아스피린, NSAIDs의 대장암 위험효과에 대해서는 아래의 체계적 문헌 고찰이 있다.

Dubé C, Rostom A, Lewin G, Tsertsvadze A, Barrowman N, Code C,

4) <http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm>

5) <http://www3.interscience.wiley.com/homepages/106568753/handbook.pdf>

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Sampson M, Moher D; U.S. Preventive Services Task Force. The use of aspirin for primary prevention of colorectal cancer: a systematic review prepared for the U.S. Preventive Services Task Force. Ann Intern Med. 2007 Mar 6;146(5):365-75.

Rostom A, Dubé C, Lewin G, Tsertsvadze A, Barrowman N, Code C, Sampson M, Moher D; U.S. Preventive Services Task Force. Nonsteroidal anti-inflammatory drugs and cyclooxygenase-2 inhibitors for primary prevention of colorectal cancer: a systematic review prepared for the U.S. Preventive Services Task Force. Ann Intern Med. 2007 Mar 6;146(5):376-89.

이 문헌의 주요 방법과 결과는 아래와 같다.

방법

- Patients : "average" risk for colorectal cancer
- * 제외 : familial adenomatous polyposis or hereditary nonpolyposis colon cancer syndromes (Lynch I or II), history of colorectal cancer
- Intervention : Aspirin, NSAIDs
- Outcome : mortality, colorectal cancer incidence, adverse events, gastrointestinal, cardiovascular, and renal harms associated with the aspirin use

결과

1) 아스피린 -benefit

Design (Studies)	Study (Reference) (Participants; Quality Rating)	Population	Dose/Duration of Regular Aspirin Use	Relative Risk (95% CI)
CRC mortality in average-risk persons				
RCT (n = 1)	Women's Health Study (32) (n = 39 876; good)	Women	Aspirin, 100 mg every other day for 10 y	NR
Cohort (n = 1)	Cancer Prevention Study II (28) (n = 1 083 531; fair)	Men	≥15 y	0.58 (0.36-0.93)
			<15 y	0.72 (0.52-0.99)
		Women	≥15 y	0.61 (0.38-0.97)
			<15 y	0.72 (0.51-1.02)

그림 6. 아스피린의 대장암 사망률 감소 효과

CRC incidence in average-risk persons				
RCTs (n = 2)	Physicians' Health Study (31) (n = 22 071; fair)	Men	Aspirin, 325 mg every other day for 5 y	1.15 (0.80-1.65)
	Women's Health Study (32) (n = 39 876; good)	Women	Aspirin, 100 mg every other day for 10 y	0.97 (0.77-1.24)
				Summary: RR, 1.02 (0.84-1.25)
Cohort studies (n = 6)	Physicians' Health Study (25) (n = 22 071; poor)	Men	Aspirin, 325 mg every other day for 12 y	1.03 (0.83-1.28)
	Health Professionals Follow-up Study (26) (n = 47 900; good)	Men	4 y	0.54 (0.34-0.83)
	Leisure World Cohort (15) (n = 13 979; poor)	Men	7-10 y	1.38 (CI not reported)
	Nurses' Health Study (30, 54) (n = 89 446; good)	Women	10 y	0.62 (0.44-0.86)
	Leisure World Cohort (15) (n = 13 979; poor)	Women	7-10 y	1.1 (CI not reported)
	North Jutland Database (29) (n = 29 470; fair)	Men and women	6 y	0.9 (0.7-1.1)
	NHANES/NHEFS (27) (n = 14 407; fair)	Men and women	NR	0.85 (0.63-1.15)
				Summary: RR, 0.78 (0.63-0.97)
Case-control studies (n = 7)	General Practice Research Database (17) (n = 12 005; good)	Men and women	>2 y	0.9 (0.8-1.1)
	Wisconsin Cancer Reporting System (23) (n = 845; fair)	Women	>5 y	0.79 (0.46-1.36)
	Juamanz et al. (21) (n = 502; fair)	Men and women	NR	0.32 (0.09-1.10)
	Multicenter Italian Case Control Study (18) (n = 3248; fair)	Men and women	5 y	0.7 (0.5-1)
	Melbourne Colorectal Cancer Study (20) (n = 1442; fair)	Men and women	NR	0.57 (0.41-0.79)
	Roswell Park Tumor Registry (19) (n = 2704; fair)	Men and women	6 y	0.33 (0.15-0.72)
	Slattery et al. (24) (n = 3051; fair)	Men and women	>5 y	0.7 (0.6-0.8)

그림 7. 아스피린의 대장암 발생률 감소

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Adenoma Incidence In average-risk persons				
RCTs (n = 1)	Physicians' Health Study (31) (n = 22 071; fair)	Men	Aspirin, 325 mg every other day for 5 y	0.86 (0.68–1.1)
Cohort studies (n = 2)	Health Professionals Follow-up Study (26) (n = 47 900; good)	Men	4 y	0.65 (0.42–1.02)
	Nurses' Health Study (30, 39) (n = 89 446; good)	Women	10 y	0.61 (0.73–0.87) Summary: RR, 0.72 (0.61–0.85)
Case-control studies (n = 5)	General Practice Research Database (35) (n = 943 903; good)	Men and women	5 y	0.9 (0.6–1.3)
	Morimoto et al. (34) (n = 1037; fair)	Men and women	3 y	0.7 (0.5–1.1)
	Logan et al. (33) (n = 476; fair)	Men and women	7 y	0.55 (0.3–1.1)
	Roswell Park Tumor Registry (19) (n = 2704; fair)	Men and women	9 y	0.61 (0.26–1.4)
	CPS-II (37) (n = 177 939; poor)	Men	10 y	0.97 (0.89–1.06)
		Women	10 y	0.85 (0.77–0.95)
				Summary: RR, 0.87 (0.77–0.98)

그림 8. 아스피린의 대장 선종 발생 감소

Adenoma Incidence/ frequency in patients with a history of colorectal adenoma				
RCTs (n = 2)	Aspirin/folate Prevention Study (42) (n = 1121; good)	Men and women	Aspirin, 81 mg/d or 325 mg/d for 1 y	0.96 (0.81–1.13)
	Association pour la prevention par l'aspirine du cancer colorectal (43) (n = 272; good)	Men and women	Aspirin, 160 mg/d or 300 mg/d for 1 y	0.61 (0.37–0.99) Summary: 0.82 (0.70–0.95)
Cohort studies (n = 2)	Polyp Prevention Study I (40) (n = 864; fair)	Men and women	"Consistent use" for 4 y	0.52 (0.31–0.89)
	Polyp Prevention Study (41) (n = 1905; good)	Men and women	Aspirin, ≤325 mg/d for 4 y	0.82 (0.65–1.02)
			Aspirin, >325 mg/d for 4 y	0.54 (0.3–0.96)
Case-control studies (n = 2)	Sandler et al. (38) (n = 492; fair)	Men and women	Aspirin, >15 tablets/mo for 5 y	0.84 (0.5–1.43)
	Breuer-Katschinski et al. (36) (n = 442; fair)	Men and women (hospital controls)	Aspirin, >4 tablets/wk for <5 y	0.91 (0.32–2.64)
			Aspirin, >4 tablets/wk for ≥5 y	0.09 (0.01–0.82)
		Population controls	Aspirin, >4 tablets/wk for <5 y	0.64 (0.26–1.56)
			Aspirin, >4 tablets/wk for ≥5 y	0.18 (0.02–1.63)

그림 9. 선종 환자에서 아스피린의 선종 예방 효과

2) 아스피린-유해

(1) 심혈관질환

Outcome	Primary Cardiovascular Prevention Outcomes (95% CI) (Reference)	Secondary Cardiovascular Prevention Outcomes (95% CI) (Reference)
All-cause mortality	No difference OR, 0.93 (0.84–1.02) (4) RR, 0.94 (0.87–1.01) (45, 57)	Reduced RR, 0.82 (0.70–0.99) (47) RR, 0.85 (0.8–0.9) (46)
Cardiovascular mortality	No difference OR, 0.87 (0.70–1.09) (4) RR, 0.93 (0.83–1.03) (45) OR, 0.89 (0.72–1.10) (57)	Reduced RR, 0.84 (0.79–0.90) (46)
Myocardial infarction	Reduced OR, 0.72 (0.60–0.87) (4) OR, 0.74 (0.68–0.82) (45) OR, 0.76 (0.67–0.85) (57) ARR, 0.5%; NNT _B , 200 (49)	Reduced RR, 0.68 (0.62–0.74) (46) RR, 0.70 (0.7–0.9) (47)
Stroke	No difference Overall: OR, 1.0 (0.85–1.23) (4); OR, 0.95 (0.84–1.06) (57) Healthy men: RR, 1.20 (0.96–1.49) (45) Cardiovascular risk factors: RR, 1.02 (0.86–1.21) (45) Hypertension: OR, 0.94 (0.76–1.17) (49)	Reduced (ischemic) Overall: RR, 0.88 (0.76–1.02) (46) and 0.8 (0.7–1.0) (47) Ischemic: RR, 0.82 (0.73–0.92) (46)
Hemorrhagic stroke	No difference OR, 1.4 (0.9–2.0) (4)	Increased RR, 1.84 (1.24–2.74) (46)

그림 10. 아스피린의 심혈관 관련 유해

(2) 위장관 질환

- 위장관 출혈 : 무작위 대조 연구의 경우 아스피린 사용군의 위장관 출혈 위험 상대 위험비는 1.6-2.5이며, 코호트 체계적 고찰인 경우 2.2, 환자-대조군 연구의 체계적 고찰은 3.1이다.
- 아스피린군의 위장관 증상의 교차비는 1.7(1.5-1.8)이다.
- 아스피린으로 인한 위장관 독성은 양-반응 관련성이 있다.

3) 아스피린의 이득, 유해 정리

(1) 이득

- 아스피린은 대장선종과 대장암 발생을 줄이는데 효과적이다. 하지만 대장암 발생 자료는 일관성이 부족(관찰 연구는 긍정적인 결과 무작위 대조 연구는 효과 없음)하고 대장암 사망 효과 감소 효과도 결과의 방향성이 서로 다르다.
- 주로 10년 이상 심혈관 질환 사용량으로 고용량으로 사용하였을 때 이득이 명백함.

(2) 유해

- 심혈관 질환 사망률, 발생률 감소 효과는 2차 예방은 비교적 명백하지만 일차 예방은 효과 없다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

- 위장관 부작용은 명백하며 주로 위장관 출혈이다. 대부분 고용량을 사용하기 때문에 용량에 따른 위험이 증가할 듯.

4) 이득과 유해의 저울질(결론)

- 아스피린은 대장 용종, 대장암 발생 위험을 줄이는 듯 하지만, 심혈관질환 예방을 위한 용량보다 더 많은 용량으로 10년 이상 비교적 오래 사용하였을 때 명백하다. 하지만 유해 가능성이 있으므로 주의가 필요하다.
- 화학예방요법의 비용-효과성에 대한 평가가 요청된다.

5) NSAIDs-benefit

Study, Year (Reference)	Participants, n	Quality Score	Population	Dose and Duration	Relative Risk (95% CI)
Effects on mortality					
Cohort study (n = 1)					
North Jutland Population Database; Lipworth et al., 2004 (21)	113 538	Fair	Average-risk men and women	Ibuprofen for $\geq 5y$	All-cause: 1.11 (1.05–1.16) [†] Bowel cancer: 0.93 (0.6–1.3) [†] Rectal cancer: 1.46 (0.9–2.3) [†]

그림 11. NSAIDs의 대장암 사망률 감소 효과

Effects on the Incidence of CRC					
Cohort studies (n = 3)					
Nurses' Health Study; Chan et al., 2005 (34)	82 911	Good	Average-risk women	≥ 2 tablets of non-ASA NSAIDs per wk for 20 y	CRC overall: 0.79 (0.64–0.97) Colon cancer: 0.71 (0.56–0.91) Rectal cancer: 1.04 (0.72–1.52)
North Jutland Population Database; Sorensen et al., 2003 (32)	183 693	Fair	Average-risk men and women	≥ 10 prescriptions for non-ASA NSAIDs over 9 y	Colon cancer: 0.7 (0.6–0.9) [‡] Rectal cancer: 0.6 (0.4–0.9) [‡]
Tennessee Medicaid Program; Smalley et al., 1999 (33)	104 217	Fair	Elderly men and women	Regular use of non-ASA NSAIDs for ≥ 1 y	0.61 (0.48–0.77)
Case-control studies					

그림 12. NSAIDs의 대장암 발생 감소 효과

RCTs (n = 3)						
PreSAP; Arber et al., 2006 (20)	933 vs. 628	Good	Higher risk (previous adenoma)	Celecoxib, 400 mg/d, for 3 y	Any adenoma: 0.64 (0.56–0.75) Advanced adenoma: 0.49 (0.33–0.73)	
APC; Bertagnoli et al., 2006 (19)	685 vs. 671 vs. 679	Good	Higher risk (previous adenoma)	Celecoxib, 400 mg/d, for 3 y Celecoxib, 800 mg/d, for 3 y	Any adenoma: 0.67 (0.59–0.77) Advanced adenoma: 0.45 (0.33–0.63) Any adenoma: 0.43 (0.31–0.61) Advanced adenoma: 0.34 (0.24–0.50)	
APPROVe; Baron et al., 2006 (18)	1158 vs. 1218	Good	Higher risk (previous adenoma)	Rofecoxib, 25 mg/d, for 3 y	Any adenoma: 0.76 (0.69–0.83) Advanced adenoma: 0.70 (0.58–0.86)	
Summary for celecoxib, 400 mg/d, or rofecoxib, 25 mg/d					Any adenoma: 0.72 (0.68–0.77) Advanced adenoma: 0.56 (0.42–0.75)	
Cohort study (n = 1)						
Polyp Prevention Study; Tangrea et al., 2002 (44)	1905	Good	Higher risk (previous adenoma)	Any NSAID use for 4 y	0.64 (0.48–0.8)	

그림 13. NSAIDs의 선종발생 감소 효과

NSAIDs - Harm

심혈관질환

Outcome	Naproxen	Non-Naproxen, Non-ASA NSAIDs	COX-2 Inhibitors
All-cause mortality	NA	No difference [52–54] to increased [21, 55] (all non-ASA NSAIDs: SMR, 1.11 (1.05–1.16) [21] and 3.4 (1.3–8.7) [55])	No difference [52]: 1.02 (0.6–1.9) vs. placebo
Cardiovascular mortality	No difference [56]: 1.47 (0.90–2.40) for COX-2 inhibitors vs. naproxen	No difference [56] (similar to COX-2): 0.67 (0.43–1.06) for COX-2 inhibitors vs. non-naproxen NSAIDs	No difference [53, 54, 56, 59]: 1.49 (0.97–2.29) vs. placebo [56] and 0.79 (0.29–2.19) [59]
Serious cardiovascular events	Neutral or reduced [56]: 0.92 (0.67–1.26) vs. placebo or 1.57 (1.21–2.03) for COX-2 inhibitors vs. naproxen	Increased [56]: 1.51 (0.96–2.37) for ibuprofen vs. placebo; 1.63 (1.12–2.37) for diclofenac vs. placebo; similar to COX-2, 0.88 (0.69–1.12) for COX-2 inhibitors vs. nonnaproxen NSAIDs	Increased [56, 58, 59]: 1.42 (1.13–1.78) vs. placebo [56]; 1.55 (0.05–2.29) [59]; 1.89 (1.03–3.45) [58]; 4.89 (1.41–16.88) for high-risk patients [58]; 1.57 (1.21–2.03) for COX-2 inhibitors vs. naproxen [56]
Myocardial infarction	Neutral or reduced [56, 59]: 0.86 (0.75–0.99) (heterogeneity) [59] or 2.04 (1.41 to 2.96) for COX-2 inhibitors vs. naproxen [56]	Increased (similar to COX-2 inhibitors) [56, 59]: 1.20 (0.85–1.68) for COX-2 inhibitors vs. nonnaproxen NSAIDs [56] and 1.55 (0.55–4.36) for rofecoxib vs. nonnaproxen NSAIDs [59]	Increased [54, 56, 58, 61]: 1.86 (1.33–2.59) for COX-2 inhibitors vs. placebo [56]; 2.04 (1.41–2.96) for COX-2 inhibitors vs. naproxen [56]; 2.93 (1.36–6.33) for rofecoxib vs. naproxen [59]; 5.0 (1.5–13.2) for rofecoxib vs. naproxen [54]; 2.83 (1.24–6.43) for high-dose rofecoxib [59]; 2.17 (1.03–4.59) for rofecoxib ≥6 mo [59]
Stroke	No difference [56, 61]: 0.08 (0.00–1.36) vs. rofecoxib [61] and 1.10 (0.73–1.65) for COX-2 inhibitors vs. naproxen [56]	No difference [56, 59]: 0.62 (0.41–0.95) for COX-2 inhibitors vs. NSAIDs [56] and 1.02 (0.54–1.93) for rofecoxib vs. NSAIDs [59]	No difference [54, 56, 58, 59, 61]: 1.02 (0.71–1.47) for COX-2 vs. placebo [56]; 1.10 (0.73–1.65) for COX-2 vs. naproxen [56]; 1.02 (0.54–1.93) for rofecoxib vs. NSAIDs [59]; 1.12 (0.43–2.91) for rofecoxib vs. NSAIDs [54]; 1.43 (0.99–2.07) for rofecoxib vs. celecoxib [58]

* Data reported are relative risks (95% CIs), unless otherwise noted. Numbers in brackets are references. NA = not applicable; SMR = standardized mortality ratio.

그림 14. NSAIDs의 심혈관 질환 유해

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

위장관

- 소화성 궤양 합병증(천공, 폐색, 출혈)은 무작위 대조 연구의 경우 OR 5.36(1.79-16.1)이며 코호트 연구의 경우 RR 2.7(2.1-3.5)이다.
- MUCOSA 연구에 의하면 소화성 궤양 합병증 비율은 연간 1.5-2.5% 정도이지만 고령이나 동반질환이 있는 경우는 연간 10%까지 상승한다.
- COX-2 inhibitors의 경우 고용량(celecoxib 200mg)에서는 위장관 사건 위험이 증가한다.

요약

- non-ASA NSAIDs는 관찰 연구에서 대장 선종 발생과 대장암 위험을 낮추며, COX-2 inhibitors 는 RCT에서 선종 환자의 이차 예방에 효과적이다.
- Non-ASA NSAIDs, COX-2 inhibitor 모두 심각한 심혈관 질환 위험을 높이며 Non-ASA NSAIDs의 위장관질환 독성 위험은 명백하며 COX-II 도 고용량에서는 위험을 증가시킨다.

결론

COX-2 inhibitor와 NSAIDs는 대장 용종 위험을 줄이고, NSAIDs는 대장암의 발생 위험도 줄인다. 하지만 이들 약제는 주요한 심혈관 질환, 위장관 질환 위해가 있다. 이득이 위험이 크다고 볼 수 없다(The balance of benefits to risk does not favor chemoprevention in average-risk individuals).

3.2.1.3 임상진료지침

임상진료지침은 "특정한 임상 상황에 적절한 관리 방법을 결정하는데 도움을 주는 체계적인 기술"로 정의할 수 있다.

임상진료지침은 몇 가지 단계로 나누어 개발한다. 뉴질랜드의 경우 1) 주제 선정, 2) 실행 가능성 선별, 3) 다학제 실무 위원회 구성, 4) 임상 질문의 확인, 5) 근거의 확인, 6) 근거평가, 7) balance sheet 작성, 8) 권고 개발, 9) 실행과 확산, 10) 갱신의 10단계로 이루어진다⁶⁾. 영국 SIGN의 경우 1) 지침 개발 조직화, 2) 주제의 선정, 3) 지침 개발 팀, 4) 체계적 문헌 고찰, 5) 지침 권고 생성, 6) 자문과 전문가 고찰, 7) 발표와 확산, 8) 실행, 9) 평가의 9단계로 이루어진다. 이중 뉴질랜드의 경우 balance

6) http://www.nzgg.org.nz/download/files/nzgg_guideline_handbook.pdf

sheet 작성과 권고 개발이 여기에 해당하고 스코틀랜드 SIGN의 경우 지침 권고 생성이 이득-위험 분석에 해당한다⁷⁾).

GRADE 접근법은 Grading of Recommendations, Assessment, Development and Evaluation (GRADE) Working Group에서 개발한 근거의 질과 권고의 강도에 대한 등급화 방법이다. GRADE 접근법은 크게 근거의 질과 권고의 강도의 등급화로 구분되며 이득-위험 분석은 권고 강도의 등급화 과정에서 하게 된다. GRADE 과정은 다음과 같은 12단계로 진행된다(Guyatt 2008).

1. 권고를 위한 특정 질문(핵심질문)을 생성한다.
2. 각 질문에 대해서 주요 결과(important outcomes)를 확인한다.
3. 결과들의 상대적 중요성을 판단 한다.
4. 근거 요약(evidence profiles)을 통해 모든 관련 근거를 요약한다.
5. 각 결과에 대해서 근거의 질을 등급화 한다.
6. 결과들 사이에서 전체 근거의 질을 결정한다.
7. 진료 옵션과 결과와 관련된 기저의 가치와 선호도에 대해서 판단한다.
8. 원하는 효과와 원하지 않는 효과에 대해 저울질 한다.
9. 이득과 비용을 저울질 한다.
10. 권고의 강도를 등급화 한다.
11. 권고를 생성한다.
12. 실행하고 평가한다.

이중 7, 8, 9, 10, 11단계가 이득-위험 분석에 해당한다.

1) GRADE 접근

GRADE 접근법에서 권고의 강도는 중재의 원하는 결과(이득)와 원하지 않는 결과(단점 - 위해, 부담, 비용) 사이의 저울질에 따라 달라진다. 이득이 단점을 초과하면 임상 의사에게 진단, 치료 전략을 적절히 선택된 환자에게 시행하도록 권고할 것이다.

만일 단점이 이득을 초과하면 해당 전략을 실행하지 않도록 권고할 것이다. 중재에 권고는 강할 수도 있고 약할 수도 있다(표 8).

7) <http://www.sign.ac.uk/guidelines/fulltext/50/index.html>

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

표 8. 원하는 결과와 원하지 않는 결과 사이의 저울질

저울질	원하는 효과가 원하지 않는 효과보다 명백히 작다	원하는 효과가 원하지 않는 효과보다 약간 작다	원하는 효과가 원하지 않는 효과보다 약간 크다.	원하는 효과가 원하지 않는 효과보다 명백히 크다.
				
권고	하지 않도록		하도록	
강도	강한	약한	약한	강한

원하는 효과에는 건강 결과의 이득(삶의 질 개선, 사건발생 감소, 입원 감소), 부담과 사망률 감소 등이 있다.

원하지 않는 효과에는 위해, 부담 증가, 비용 증가가 있다. 부담은 권고대로 시행하는데 환자나 보호자에게 필요하지만 좋아하지 않는 것으로 예들 들어 약제 투약, 병원 방문의 불편함 등이다.

원하는, 원하지 않는 효과에 대한 저울질을 통해 권고를 할 때는 다섯 가지 요인(결과의 중요성, 결과의 기저 상태 위험, 중재의 절대, 상대 효과, 효과 추정의 정밀도, 비용)을 고려하여 결정한다.

첫째, 결과의 중요성이다.

일반적으로 이로운 결과의 가능성을 높이는 중재 혹은 이상 결과의 위험을 줄이는 중재에 대해 강한 권고를 하게 된다. 권고를 하기 전에 결과의 중요성을 고려하여야 하며, 환자가 결과에 대해 가지는 실제적, 상대적 가치를 찾아야 한다. 또한 환자나 다른 이해 당사자의 선호도와 가치의 변이성에 대한 근거도 찾아야 한다. 가치나 선호도의 변이가 크면 강한 권고를 받기 힘들다.

둘째, 기저 상태 위험이다. 일반적으로 기저 상태 위험이 높으면 이득의 효과가 더 커지고 권고의 강도가 높아진다. 기저상태가 의미있게 이득이나 위해의 크기에 영향을 미치고, 결과 발생 위험에 대한 특정 인구 집단이 존재하면 권고를 이들 다른 인구 집단에 따라서 구별하여 하는 것을 고려하여야 한다.

셋째, 절대, 상대 효과와 권고의 강도이다. 만일 중재의 효과들이 일관성있게 한쪽 방향(이득이건 단점이든)을 향한다면 이 경우 강한 권고 등급을 받을 가능성이 크다. 만일 중재의 효과들이 다른 방향을 향한다면(이득이 크지만 부작용도 크다) 약한 권고 등급을 받을 가능성이 크다.

넷째, 효과 추정치의 정밀도와 권고의 강도이다.

효과 추정치가 정밀하면 더 강한 권고 등급을 받을 가능성이 커진다.

다섯째, 가치와 선호도에 대한 확신이다.

가치와 선호도에 대한 불확실성과 변이가 있으면 권고의 강도가 낮아진다. 서로 다른 전략에는 장점과 단점이 있으며 환자는 그 둘 사이에서 저울질(trade-off)을 하여야 한다. 따라서 지침패널이 어떻게 특정 이득, 위험, 불편함에 대해 가치를 부여하는지가 권고와 권고의 강도를 결정하는데 중요하다. 서로 다른 결과에 대한 가치와 선호도에 대한 연구는 거의 없기 때문에 가치와 선호도에는 언제나 불확실성이 있다고 주장한다. 반면에 가치와 선호도에 대한 몇몇 체계적 연구가 있으며 지침패널의 환자 진료의 경험이 추가적인 통찰력을 제공한다. 가치와 선호도에 대해 환자에게 직접 정보를 얻는 것이 가장 이상적이지만 연구를 통해 자료를 얻을 수도 있다. 일반적으로 이런 연구는 이용 가능하지 않다. 권고의 해석에 가치와 선호도에 대한 판단이 매우 중요하면 권고를 할 때 중요하게 생각한 가치에 대해 기술하여야 한다.

임상진료지침 시나리오

다양한 위험도를 가진 군에게 대장암 혹은 대장 선종의 예방을 위해 아스피린 혹은 NSAIDs를 투여하는 것이 필요한지에 대하여 권고안을 만들기로 하였다. 해당 주제에 대한 체계적 문헌 고찰을 검토하여 권고안을 만들었다

이에 대한 USPSTF의 분석 결과는 아래 아래와 같다.

이득 관련

1. 아스피린과 NSAIDs를 고용량으로 장기간 사용하면 대장 선종의 발생을 줄인다는 중등도 이상의 근거가 있다.
2. 저용량 아스피린은 대장암 발생을 줄이지 못한다는 높은 근거가 있다.
3. 심혈관 질환 예방때 사용 용량보다 높은 수준의 아스피린과 NSAIDs는 대장암 발생을 줄인다는 중등도 근거가 있다.
4. 장기간 동안 아스피린을 사용하면 대장암 발생을 줄인다는 중등도의 근거가 있다.
5. 아스피린과 NSAIDs를 사용하면 대장암 사망률을 낮춘다는 낮은 수준의 근거가 있다.

유해관련

1. 아스피린은 위장관 출혈 위험을 양-반응 관계로 높인다는 높은 수준의 근거가 있다. 아

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

- 스피린은 출혈성 뇌졸중 위험을 높인다는 높은 수준의 근거가 있다.
2. NSAIDs는 위장관 출혈과 신장 손상을 높인다는 높은 수준의 근거가 있다.
 3. COX-II inhibitor는 신장 위험을 높이고 심혈관 질환의 위험을 높인다는 높은 수준의 근거가 있다.
 4. 전체적으로 아스피린과 NSAIDs는 중등도의 유해가 있다는 높은 수준의 근거가 있다.
- 이와 관련한 권고 사항은 아래와 같다.

대장암의 평균적 위험을 가진 성인에 있어서 대장암 예방을 위해 아스피린 혹은 NSAIDs를 복용하는 것은 권고되지 않는다(grade D)

3.2.2 여러 기구에서 BRA

전 세계적으로 이득-위험 분석을 가장 많이 사용하고 있는 기구는 약제 허가 기구라고 할 수 있다. 약제 허가 기구는 EU의 경우 EMEA(European Medicines Agency), 미국 Food and Drug Administration (FDA), 영국 MHRA (Medicines and Healthcare products Regulatory Agency), 호주 TGA (Therapeutic Goods Administration), 캐나다 Health Products and Food Branch (HPFB) of Health Canada 등이다. 이중 실제로 이득-위험 분석에 대한 논의를 진행하고 있는 대표적인 기구는 EMEA와 FDA이다.

3.2.2.1 EMEA

EMEA는 유럽 연합에서 통합적으로 운영하고 있는 약제 승인 기구이다. EMEA에서 승인을 받으면 다른 나라에서도 인정을 받는 제도를 갖고 있다(centralized). 현재 25개 EU 국가에서 EMEA를 받아들이고 있으며, 42개 센터 과학 시설 같이 이용하고 있다. EMEA 기구로는 사람의 약제에 대한 심의 기구인 Committee for Medicinal Products for Human Use(CHMP), 가축 약제 심의 기구인 Committee for Medicinal Products for Veterinary Use(CVMP, 가축), 희귀 약품에 대한 기구인 Committee on Orphan Medicinal Products(COMP), 허브 관련 심의 기구인 Committee on Herbal Medicinal Products(HMPC)로 나눌 수 있다(박병주 2009).

EMEA에서는 2008년 3월 "REFLECTION PAPER ON BENEFIT-RISK ASSESSMENT METHODS IN THE CONTEXT OF THE EVALUATION OF

MARKETING AUTHORISATION APPLICATIONS OF MEDICINAL PRODUCTS FOR HUMAN USE"라는 제목의 보고서를 채택하였다. 이 보고서는 2007년 1월 초안이 나온 이후 각계각층의 의견을 수용하여 최종안이 나온 것이다. 이 보고서는 신약의 이득과 위험을 평가함에 있어서 이득-위험 평가에서 1) 방법론, 2) 일관성과 투명성을 향상시킬 수 있는 권고안을 제시하는 것이 목적이다. 현재 이 보고서가 파악한 전체적인 상황은 1) 여전히 전문가 판단이 이득-위험 평가의 가장 중요한 부분으로 남을 것이며, 2) 양적인 위험-이득 분석이 질적인 평가를 대체하지는 않을 것이라는 것이다.

그럼에도 관심을 기울여야 할 이득-위험 평가 방법은 1) 가장 중요한 이득과 가장 심각한 위험을 쉽게 확인하게 해주는 것, 2) 개별 이득과 위험에 대한 명백한 가중치를 부여하는 것, 3) 근거의 강도와 확인된 불확실성에 대해 제시하는 것이다. 이러한 분석에 기초해서 CHMP는 추후 두 가지 단계를 거쳐서 이러한 방법론을 고려하기로 하였는데 첫 번째 단계는 현재 CHMP assessment report templates 중 위험-이득 평가 부분을 개선해서 구조화된 이득, 위험 기준과 지침을 통합하여 제시하도록 개선하는 것이다. 개선되는 templates는 첫째 구조화되고 양적인 접근을 하는 것, 둘째, 특정 치료 상황에서 이득과 위험성의 중요성에 대한 명백한 정의를 하는 것, 셋째, 불확실성과 변이성의 원천, 이득-위험 평가에 미치는 영향 등에 대하여 분석을 하는 것이다.

두 번째 단계는 추후 연구 활동과 연대 활동을 하는 것으로 우선 추후 비용/이득 분석 방법론 발전(quantitative and semi-quantitative) 상황을 주시하여 추후 연구 사업 시행하며, 비용/이득 방법론 관련 국제단체, 유럽 내 기구와 연대 활동 하는 것이다(CHMP 2008).

3.2.2.2 FDA

FDA는 약제안전과 위험관리 부서(Drug Safety and Risk Management division)를 설립하였는데 이 부서에서 안정성 평가, 효능, 약제의 남용 가능성과 함께 위험 관리, 위험 교류와 자발적 보고 등을 관장한다. 이런 분석을 어떻게 하는지에 대해서는 잘 알려져 있지 않다. 자문 기구 등 여러 가지 위원회가 존재한다.

미국 FDA는 현재의 상황을 다음과 같은 6가지로 요약하였다.

첫째, 체계적이고, 투명한 이득-위험 분석이 부재하다. 현재까지 여러 가지 이득-위험 분석이 제안되었지만 충분히 체계적이고 전체 내용을 충분히 전달할만한 방법론이 현재까지는 없다는 것이다.

둘째, 이득-위험 분석의 핵심은 이득과 위험의 저울질이다. 둘 사이에 어떤 것이 더 큰지에 따라서 의사 결정을 하면 된다. 하지만 현재까지 제안된 어떠한 방법을 사용하여도

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

이득과 위험 저울질에 있는 불확실성을 없앨 수 없다. 방법에 따라 결과의 일관성이 없으며 어떤 방법론을 사용하는 것이 가장 진실에 근접한 방법론인지에 대한 합의가 없다.

셋째, 허가 당시 약제 위험에 대한 지식이 불충분하다. 이득-위험 분석이 가장 많이 요청되는 경우는 약제 허가 당시이다. 하지만 약제를 허가할 당시에는 약제의 효능에 대한 임상시험 정보만 알려져 있기 때문에 약제 허가 이후 알게 되는 다양한 상황이나 인구 집단에서의 효과나 비용 정보 그리고 약제의 이상 사건이나 반응 등에 대한 자세한 정보를 알기 어렵기 때문에 적절한 이득-위험 분석을 하기 어렵다.

넷째, 이득과 위험 평가에서 이해관계(conflict of interest)가 문제가 될 수 있다.

다섯째, 약제의 이득-위험 분석에는 다양한 이해 당사자가 참여해야 하지만 현재까지의 방법론에서는 약제에 대한 의사 결정의 주요 이해 당사자라고 할 수 있는 임상 의사와 공공부분에서 참여할 수 있는 기재가 없다.

여섯째, 이득-위험 분석에 사용되는 여러 가지 용어는 혼란스럽고 비일관적이다. 이득과 위험은 각각 무엇인지에 대해서 방법론마다 달리 정의하며 개별 방법론에서도 연구에 따라서 변이가 존재한다⁸⁾.

이를 바탕으로 새로운 이득-위험 분석의 조건으로 다음과 같은 것을 제시하였다(CHMP 2008).

- 1) 약제 이득 위험 분석의 복잡성을 잘 포착하여야 한다.
- 2) 개별 환자에서 이득-비용의 특성을 양적으로 표현할 수 있어야 한다.
- 3) 새로운 정보가 나올 때 갱신할 수 있어야 한다.
- 4) 이득-위험 평가에 내재한 불확실성을 다루어야 한다.
- 5) 이득-위험 평가에서 비용의 문제를 해결하여야 한다.
- 6) 양적 체계 도입의 비용 문제와 잠재적 이상 효과에 대하여 다루어야 한다.
- 7) 양적인 정보를 효과적으로 제시하고 교류할 수 있게 하여야 한다.

- 현재 FDA에서는 QALY-Based Approach to Benefit-Risk Modeling을 기본적인 방법론으로 채택하는 것을 고려중이다(Incremental Net Health Benefit, INHB).

8) Institute of medicine. Understanding the benefits and risks of pharmaceuticals. workshop summary. the National academies press. 2007.

3.2.3 주요 위험-이득분석

EMA의 보고서는 이득-위험 분석의 방법론을 다음과 같이 구분한다⁸⁾.

1. Number Needed to Treat/Number Needed to Harm (NNT/NNH)에 기초한 방법론
2. quality-adjusted life years (QALYs)에 기초한 방법론
3. multi criteria decision analysis (MCDA)
4. “Principle of three” and the TURBO methods
5. large body of research developed by pharmaco-epidemiologist

3.2.3.1 Number Needed to Treat/Number Needed to Harm (NNT/NNH) method

NNT나 NNH를 이용한 이득-위험분석에는 여러 가지가 있지만 대표적인 것으로 Likelihood of being helped or harmed(LHH)(McAlister 2000), Threshold NNT(NNTt)(Sinclair 2001), Relative-Value Adjusted Number-Needed-To-Treat (RV NNT)(Holden 2003) 등이 있다.

1) Likelihood of being helped or harmed(LHH)

비록 NNT와 NNH가 임상 의사들이 치료의 이득-손실을 환자에서 설명할 때 편리한 면은 있지만, 각 사건의 개별 위험도를 모두 알고 싶어하는 환자에게는 정보의 내용이 호소력이 없을 수 있다. 환자의 가치판단을 통합하여 정보를 제공하는 방법들이 최근에 개발되고 있는데, 이것이 LHH이다.

이 방법의 첫 번째 단계는, 치료를 받는데(혹은 받지 않는데) 대한 환자의 가치를 탐색하고, 치료에 따른 부작용의 정도(혹은 치료를 통해 회피하려고 하는 부정적 사건의 심한 정도)에 대한 가치를 조사하는 것이다. 환자에게 0(사망)에서 1(완전한 건강)까지 표시된 등급척도를 제공하고 부정적 사건의 가치를 체크해 보도록 한다.

다음의 예를 생각해 보자

응급실을 통해서 뇌졸중 환자 2명이(환자 A는 65세 여자; 환자 B는 65세 남자)입원하였다. 신체검사 상 두 사람 모두 우측 상지의 가벼운 쇠약감과 좌측 경동맥의 잡음이 있었다. 환자 A는 고혈압의 과거력이 있었고 입원시 혈압은 200/110이었다. 기타 다른 이상은 없었다. 경동맥 잡음에 대해 도플러 초음파를 시행하였다. 두 환자 모두 혈관 내피에 불규칙한 병변이나 궤양은 없이 약 50-69% 정도의 중등도 협착만 초음파상 관찰되었다. 증상이 있는 경동맥 협착에 대해 내과적 치료와 외과적 치료 중 어떤 것을 적용하는 것이 타당한지 문헌을 검토해 보는 것이 좋겠다고 생각하였다. 경동맥 내막절제술과 표준 내과적 치료(아스피린)를 비교하는 무작위 대조군 연구들을 체계적으로 검토하기 위해 인터넷 "pubmed"에서 "stroke[mh] AND carotid endarterectomy[mh] AND randomized clinical trial[PT]"로 검색하였고 가장 근접한 논문으로 한 논문(Barnett HJ, Taylor DW, Eliasziw M, et al. Benefit of carotid endarterectomy in patients with symptomatic moderate or severe stenosis. *N Engl J Med*. 1998;339:1415-1425)을 찾았다.

앞의 예에서 환자 A는 뇌졸중에 의한 장애를 즉각적인 사망과 마찬가지로 생각해서 그 가치를 0으로 체크하고 치료에 따라 초래될 수 있는 부작용(수술 30일 이내 사망)에 대해서는 0.25를 체크했다고 하자. 이 환자는 장애를 남길 뇌졸중이 다음 달에 사망하는 것보다 1.3배 나쁘다고 믿는다고 할 수 있다 $[(1-0)/(1-0.25)]$. 반면에, 환자 B는 사망하는 것이 장애가 남는 뇌졸중보다 5배나 좋지 않다고 생각한다고 하자.

중등도의 경동맥 협착증을 가진 뇌졸중환자는 5년 동안 뇌졸중이 생길 가능성은 10.3% 정도이고 내막 절제술을 행하면 이 수치를 5.3%까지 줄일 수 있다. 이런 환자의 평균 NNT는 20이다. 내막 절제술을 시행한 경우 사망의 절대 위험 증가는 1.6%이고 평균 NNH는 63(1/0.02)이다. 당신병원의 신경(혈관)외과의의 수술전후 사망률이 2%라면 이 연구의 NNH를 당신의 환자에게 적용해 볼 수 있을 것이다.

LHH는 1/NNT 즉 절대 위험 감소와 1/NNH 즉 절대 위험도 증가의 비이다. 이 경우 $LHH = (1/NNT) : (1/NNH) = (1/20) : (1/63) = 3$ 대(對) 1로 수술을 하는 것이 더 낫다. 이 수치를 사용해서 환자에게 “수술이 당신에게 끼칠 손해보다는 이득이 3배나 높습니다.”라고 이야기 할 수 있다. 그러나, 이 수치는 각 개인의 개별적인 위험도, 뇌졸중과 수술후 사망 등의 대한 가치 판단을 고려에 넣지 않은 것이다. f 요인을 이용하여 각 환자들의 LHH를 구할 수 있다. 위에서 언급한 대로 여자는 뇌졸중의 위험이 상대적으로 낮으므로, 환자 A에게는 ft를 대략 0.2로 계산할 수 있다. 또한 여러 논문에서 여자환자, 좌측 경동맥 협착증, 고혈압의 기왕력 등이 수술전후 사망률을 높였다(RRs=1.4-2.3)(Barnett 1998). 따라서, 환자 A의 경우 수술로부터 사망할 위험이

커진다($f_h=2$).

이 경우 위험도를 보정한 LHH는 $LHHA = [(1/NNT) \times f_t : [(1/NNH) \times f_h]$
 $= [(1/20) \times 0.2] : [(1/63) \times 2] = 3 \text{ 대(對)} 1 \text{로 내과적 치료가 더 낫다.}$

비슷하게 환자 B의 경우도 하위그룹으로부터 계산한 환자 B의 f_t 는 대략 1.25가 된다.
 환자 B 또한 좌측 협착증이므로, 수술전후의 이 환자의 사망 위험도는 증가한다($f_h=2$).
 위험도를 보정한 LHH는 $LHHB = [(1/20) \times 1.25] : [(1/63) \times 2] = 2 \text{ 대(對)} 1 \text{로 수술이 낫다고 나타난다.}$

위험도를 보정한 LHH는 여전히 각 환자의 가치판단을 포함하지 않은 것이다. 환자 A는 사망하는 것보다 장애가 남는 것을 1.3배 두려워하므로는 LHH(s = severity factor)는 $LHHA = [(1/NNT) \times f_t \times s] : [(1/NNH) \times f_h] = [(1/20) \times 0.2 \times 1.3] : [(1/63) \times 2] = 2 \text{ 대(對)} 1 \text{로 내과적 치료가 더 낫다.}$ 이리하여 환자 A의 가치판단이 포함된 개별화된 이득-손실 평가에서, 내과적 치료가 2 배정도 이득이 될 것이라고 예측할 수 있다. 반면에, 환자 B는 사망하는 것이 장애가 남는 것보다 5 배정도 나쁘다고 생각하므로 이것을 통합하여 계산하면 $LHHB = [(1/20) \times 1.25] : [(1/63) \times 2 \times 5] = 3 \text{ 대(對)} 1 \text{로 내과적 치료가 더 낫다.}$

이 두 가지 예는 어떻게 환자 개개인의 가치판단을 임상 결정 과정에 참여시키는지를 보여준다. 현재 이 과정은 시간적 소모가 많고, 부정확한 면이 있으며, 이에 따라 환자의 결과가 얼마나 달라질지 검증되진 않았지만 컴퓨터화 된다면 임상적으로 더 유용해 질 것이다. 만약 환자의 f 값에 대한 확신이 없거나, 환자의 가치판단에 대한 평가가 확실치 않다면, 이런 변수들에 다른 값을 대입했을 때 LHH에 미치는 영향을 보기 위해 민감성 분석을 해 보아야 한다. 우리는 경동맥 내막절제술과 관련된 다른 요소나 진단 검사의 위험도 등을 고려하지 않은 단순화된 공식을 사용하였지만 더 복잡한 상황에는 수정하여 사용할 수 있다.

2) Threshold NNT (NNTt)(Sinclair 2003)

특정 치료법을 시행하려면 해당 치료의 이상 반응과 비용보다 치료 효과를 비교하여 치료의 효과가 더 커야 한다. 만일 이상반응과 치료 효과가 더 크다면 치료법을 선택해서는 안된다. 치료 효과를 NNT로 계산하였을 때 해당 요법의 NNT는 치료효과의 치료의 위험이 같아지는 NNT 역치(Threshold NNT, NNTt)보다 커야 한다. NNT 값은 목표 사건(target event)이 발생할 위험에 따라 달라지기 때문에 NNTt 값은 목표 사건의 특정 발생 위험을 가지는 특정 환자 집단에 따라서 달라지며 이 집단에서 NNTt 값을 결정하여야 한다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

NNTt 값을 비용을 포함하는 경우(full model)과 포함하지 않는 경우로 구분한다.

① 비용을 포함하는 경우

NNTt 값을 구하기 위해서는 다음과 같은 자료가 필요하다.

- 한 건의 목표 사건 회피를 위한 치료 비용
- 개별 환자 당 해당 치료비용 (a) 치료의 직접 비용 (b) 치료 관련 이상 반응 치료비용
- 목표 사건 회피의 가치(목표 사건 예방을 위해 지불할 용의가 있는 가치, 재화 비용)
- 치료 관련 부작용(들)의 가치(이상 반응(들)을 예방하기 위해 지불할 용의가 있는 재화 비용)
- 치료 관련 이상 반응 비율

이들은 다음과 같은 약자로 표현된다.

비용

- TEC(Target event cost) 목표 사건 비용
- AEC(Adverse event cost) 이상 반응 비용
- DC (Direct cost of treating a patient) 한 환자 치료 직접 비용

가치(Values) 관련

- TEV(Target event value) 목표 사건 가치
- AEV (Adverse event value) 이상 반응 가치

발생률

- TER(Target event rate) 목표 사건 발생률
- AER(Adverse event rate) 목표 사건 발생률

NNTt 값에서 치료 비용과 가치

비용

- TEC 한 건의 목표 사건 예방을 통해 절약할 수 있는 비용
- DC·NNTt NNTt 명의 환자 치료 비용
- AEC·AER· NNTt NNTt 명의 환자를 치료할 때 예상되는 이상 반응 치료 비용

가치

- TEV 한 건의 목표 사건 예방의 가치
- AEV·AER·NNTt NNTt 명의 환자를 치료할 때 예상되는 이상 반응으로 인한 가치의 감소

NNTt는 치료 관련 실제 비용(net cost)과 치료 관련 실제 가치(동일한 경제학적 가치로 환산, net value)가 같다는 공식에서 얻는다.

이는 아래와 같은 식으로 표현할 수 있다.

$$TEC + (DC \cdot NNTt) + (AEC \cdot AER \cdot NNTt) = TEV \cdot AER \cdot NNTt$$

이 식에서 NNTt를 얻을 수 있다.

$$NNT_T = \frac{TEV + TEC}{DC + AER(AEV + AEC)}$$

만일 이상 반응이 여러 개 이면 아래와 같은 식에서 얻는다.

$$NNT_T = \frac{TEV + TEC}{[DC + AER_1 \times (AEV_1 + AEC_1) + \dots AER_n (AEV_n + AEC_n)]}$$

② 비용을 제외한 경우

치료의 경제학적 비용을 고려하는 것이 적절하지 않으면 목표 사건과 이상 사건에 대한 가치만으로 NNTt를 구할 수 있다. 이 경우 아래와 같은 식으로 표현할 수 있다.

$$NNT_T = \frac{TEV}{AER_1 \times AEV_1 + \dots AER_n \times AEV_n}$$

이 경우 절대 가치를 알아야 하므로 상대적 가치를 알고 있다고 하면

$$NNT_T = \frac{1}{AER_1 \times RV_n + \dots AER_n \times RV_n}$$

여기에서 RV는 목표 사건에 대한 이상 반응의 상대적 가치를 말한다.

치료가 정당화 되는 최소 사건율(minimum event rate for treatment, MERT).

무작위 대조연구에서 대조군의 목표 사건의 위험 수준은 대조군 사건 발생률로 표시한다. 대조군의 여러 하부 집단에 따라서 위험도는 평균보다 크기도 작기도 하지만 상대적 위험 감소는 위험군 수준에 따라서 대부분 일정하다. 상대 위험도 감소가 일정하다고 가정할 때 NNT는 아래와 같이 표현된다. CER은 대조군 사건 발생률(control event rate)이고 RRR은 상대 위험도 감소(relative risk reduction)이다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

$$NNT = \frac{1}{CER \times RRR}$$

이때 NNT가 NNT_t일 때를 결정적 CER(critical CER)이라고 한다(표 8에서 a). 이 때 CER을 치료의 최소 사건 발생률(minimum event rate for treatment, MERT)이라고 한다. 따라서 아래와 같은 수식이 성립한다.

$$NNT_T = \frac{1}{MERT \times RRR}$$

$$MERT = \frac{1}{NNT_T \times RRR}$$

아래 [그림 15]에서 D영역의 경우 해당 CER에서 NNT값이 NNT_t보다 크기 때문에 치료하지 않는 것을 권고하게 되고 A영역의 경우 해당 CER에서 NNT값이 NNT_t보다 작고 95% 신뢰구간을 벗어나기 때문에 권고할 수 있다.

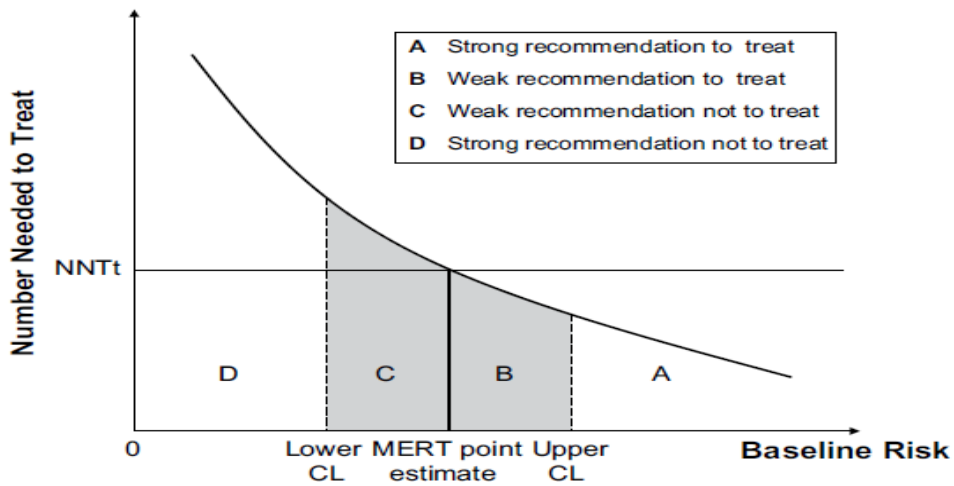


그림 15. NNT threshold, target event에 대한 환자의 기저 위험도, NNT간의 관계

예시)

심방세동 환자에서 혈전 색전의 위험은 나이, 뇌졸중 과거력, 고혈압 병력, 당뇨병의 네 가지 위험인자에 의해 결정된다. 메타분석 결과 와파린은 뇌졸중 발생 위험을 68%(50-79%) 줄인다. 또한 와파린이 색전성 뇌졸중을 줄이는 절대 효과(ARR)는 CER이 연간 0.5-8.0%로 변화하면서 0.0034-0.0544(NNT 18-294)로 변화한다(표 8).

이 경우 RRR은 모두 일정하다고 가정한다. 이상사건 중 심한 출혈의 경우 두 군간 차이는 0.6%이다. 심한 출혈 환자 중 20%는 사망하고, 3%는 심하지만 사망에 이르지 않는 뇌출혈이고, 8%는 경미한 뇌출혈이며, 69% 심하지만 비치명적인 위장관 출혈이다. 비용 자료에 의하면 심한 뇌출혈, 경미한 출혈, 심한 위장관 출혈은 각각 \$34,200, \$7,800, \$3,920의 비용이 들며 1년간 항응고제 비용은 \$800이 필요하다.

가치를 추정하기 위해서는 많은 가정이 필요하다. 일단 효용상실 기간을 경미한 뇌졸중, 주요 위장관 출혈, 경미한 출혈은 1년으로 하고 심각한 뇌졸중과 사망은 영구(여기서는 10년)로 가정한다. 가치 판단은 영구적인 상태 중 사망, 심각한 뇌졸중, 심각한 뇌출혈은 각각 3, 5, 5건이 동급이다. 또한 효용상실 기간이 1년이 상태 중에서 경미한 뇌졸중과 뇌출혈, 심각한 위장관 출혈, 경미한 출혈은 각각 15, 7.5, 400건이 동급이다. 이러한 자료를 바탕으로 완전한 건강으로 1년 생존하는 것의 효용성을 1.0로 했을 때, 뇌졸중으로 사망하는 것은 0, 심각한 뇌졸중은 0.4, 경미한 뇌졸중은 0.8, 주요 위장관 출혈은 0.6, 경미한 출혈은 0.993이라고 할 수 있다.

NNTt를 계산하기 위한 기초 자료는 아래와 같다.

결과	효용성 상실	기간(년)	총상실(QALYs)\$	비용의 현재 가치
치명적	1.0	10	10	0
심한 뇌졸중	0.6	10	6	34,200
경미한 뇌졸중	0.2	1	0.2	7,800
위장관 출혈	0.4	1	0.4	3,920
경미한 출혈	0.007	1	0.007	100

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

목표 사건 : 뇌졸중

뇌졸중 종류	상대빈도	비용의 현재 가치		효용성 상실	
		비용	비용 × 상대빈도	QALYs	QALYs × 상대빈도
치명적	0.25	0	0	10	2.5
심한 뇌졸중	0.25	34,200	8,550	6	1.5
경미한 뇌졸중	0.5	7,800	3,900	0.2	0.1
모든 뇌졸중			12,450		4.1

이상 사건 : 출혈

출혈 종류	상대 빈도	기여 위험도	비용의 현재 가치		효용성 상실	
			비용	비용 × 상대빈도	QALYs	QALYs × 상대빈도
심각한 출혈						
치명적	0.20	0.0012	0	0	10	2
심한	0.03	0.00018	34,200	1,026	6	0.18
경미한	0.08	0.00048	7,800	624	0.2	0.016
위장관	0.69	0.00414	3,920	2,705	0.4	0.278
AE:모든 심각		0.0016		4,355		2.472
AE:경미한 출혈		0.15	100	12,450	0.007	

사건의 효용성 상실의 상대 가치

결과	효용성 상실(QALYs)	효용성 상실의 현재 가치
목표 사건 : 모든 뇌졸중	4.1	1.0
모든 심각한 출혈	2.472	0.603
모든 경미한 출혈	0.007	0.00171

Threshold NNT는

$$\begin{aligned}
 NNT_T &= \frac{TEV + TEC}{[DC + AER_1 \times (AEV_1 + AEC_1) + \dots AER_2 (AEV_2 + AEC_2)]} \\
 &= \frac{205,000 + 12,450}{[800 + 0.006 \times (123,600 + 4355) + 0.15 (350 + 100)]} \\
 &= 133
 \end{aligned}$$

비용을 고려하지 않은 NNT는

$$\begin{aligned} NNT_T &= \frac{1}{AER_1 \times RV_1 + AER_2 \times RV_2} \\ &= \frac{1}{0.006(0.603) + 0.15(0.00171)} \\ &= 258 \end{aligned}$$

아래의 식으로 MERT를 계산하면

$$MERT = \frac{1}{NNT_T \times RRR} = 0.011 (\text{비용 계산하지 않으면 } 0.006)$$

위의 예에서 75세로 위험인자가 1개 있으면 CER은 8.1%로 MERT인 1.1%보다 명백히 크므로 항응고치료를 하는 것이 맞으며 65세로 위험인자가 없는 경우 CER은 1.0%로 1.1%와 유사하므로 어느 쪽을 선택하더라도 위험과 이득은 유사하다고 할 수 있다.

3) Relative-Value Adjusted Number-Needed-To-Treat (RV NNT)

NNT라는 개념이 새로운 것은 아니지만 이득-위험 분석에 사용하는 것은 비교적 최근 개념가치이다. NNT는 치료로 인한 절대 위험 감소의 역수이며 하나의 부가적인 악결과 예방하기 위하여 치료하여야 하는 환자의 수로 정의할 수 있다. NNT는 아래의 식으로 계산할 수 있다(p1은 대조군에서 질병 비율, p2는 치료군에서 질병 확률).

$$NNT = \frac{1}{p1 - p2}$$

마찬가지로 number-needed-to-harm (NNH)는 아래와 같이 계산할 수 있다(q1과 q2는 각각 치료하지 않는 군과 치료 받은 군에서 나타나는 이상 반응 비율이다).

$$NNH = \frac{1}{q2 - q1}$$

치료는 NNT가 위험보다 높은 것이 목표이므로 $NNT/NNH < 1$ 혹은 $NNT < NNH$ 가 목표이다. 이 식은 추가적인 질병 발생을 위해 치료해야 하는 환자의 수가 한 사람의 부가적인 이상반응을 예방하기 위해 치료해야하는 수보다 적어야 한다는 의미이다. 이런 방

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

법의 문제점은 질병과 이상 반응의 중요도가 고려되지 않은 것이다. 이런 문제점을 해소하기 위하여 질병과 이상 반응의 환자의 상대적 가치(Relative value)를 고려하는 것이며 여기에서 가치란 특정 결과에 대한 환자 선호도이며 아래와 같이 표시한다.

$$RV = \frac{1 - \text{이상반응의 효용성}}{1 - \text{질병의 효용성}}$$

이를 반영하여 NNH를 아래와 같이 표시할 수 있다.

$$NNH = \frac{1}{(q2 - q1) \times RV}$$

만일 이상 반응의 수가 여러 개이면 아래와 같이 표시할 수 있다.

$$NNH_{RV-adjusted} = \frac{1}{(q21 - q11) \times RV_1 + (q22 - q12) \times RV_2 + \dots (q2k - q1k) \times RV_k}$$

다음의 예를 생각해 보자

류마티스 관절염 치료의 표준 약제인 Methotrexate(MTX)와 새로운 약제인 Drug A의 NNT 과 $NNH_{RV-adjusted}$ 값은 아래 표와 같다

아래 표에서 E1은 염증 반응이 20% 호전되는 점(American College of Rheumatology, ACR 20)이고 E2는 50%(ACR 50), E3는 70%(ACR 70)호전되는 점이다.

표 9. 류마티스 관절염 치료 약제간의 RV NNT

		Drug A - placebo	MTX - placebo	Drug A - MTX
NNT		4	5	15
NNH _{RV-adjusted}	E1 utility = 0.40	42	70	108
	E1 utility = 0.23	24	40	62
	E1 utility = 0.10	11	17	27
NNT		4	7	9
NNH _{RV-adjusted}	E2 utility = 0.40	42	70	108
	E2 utility = 0.23	24	40	62
	E2 utility = 0.10	22	17	27
NNT		6	19	9
NNH _{RV-adjusted}	E3 utility = 0.40	42	70	108
	E3 utility = 0.23	24	40	62
	E3 utility = 0.10	11	17	27

위의 표에서 신약 A는 효과와 안전성을 비교하였을 때 MTX에 비해 더 우월하다고 할 수 있으며 이는 ACR 20, 50, 70 모두에서 그러하다. 즉 $NNT < NNH_{RV-adjusted}$ 가 E1, E2, E3에 모두 해당한다고 할 수 있다. 이를 E1 즉 ACR20으로 설명하면 NNT값이 15는 NNH_{RV-adjusted} 값이 108, 62, 27 보다 우월하기 때문에 이상반응 발생을 위해 필요한 환자수보다 효과를 얻기 위해 치료해야 하는 환자수가 더 적은 것이다.

이 방법의 단점은 두 가지이다. 첫째, NNT가 절대 위험감소의 역수이다. 역수의 경우 몇 가지 문제점이 생기는데 특히 ARR이 0인 경우 해석이 힘들다는 것이다. 두 번째는 효용성(utility)의 이용 가능성이다. 일부 선호도에 대한 자료가 있기는 하지만 해당 분야에 대해 연구가 진행되지 않았으면 자료를 얻기 힘든 문제가 있다.

3.2.3.2 Quality-adjusted life years (QALYs)

Quality-adjusted life years는 의학적 중재의 이득을 정량화 하는 수단으로 살아가는 기간과 질을 같이 고려한다. 대표적인 것으로 Q-TWiST(A quality-adjusted time without symptoms or toxicity) method(Staquet 1998)와 Incremental Net Health Benefit(INHB)가 있다(McAlister 2000).

현재 이득-위험 분석에서 QALYs의 역할을 불분명하다. 분명히 QALY가 건강 결과에 대한 편리하고, 단순한 지표이기 때문에 이득과 위험 사이의 저울질 수단으로 적절하기는 하다. 하지만 이득-위험의 단일 지표로서 QALYs는 윤리적, 개념적, 조작적 차원에서 비

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

판을 받아왔다. 또한 여명이 길 것으로 예상되는 경우 조기 사망이나 삶의 질 향상 등을 적절히 반영하지 못하며 비율이 아니기 때문에 생길 수 있는 오류 가능성, 효용성과 시간을 하나의 지표로 만든 것의 문제 등이 지적되고 있다(Felli 2009).

1) Q-TWiST(A quality-adjusted time without symptoms or toxicity) method

Gelber 등은 질병 증상과 독성 효과가 없는 질 보정 생년(quality of life adjusted time without symptoms of disease, TWiST and Q-TWiST)의 방법을 개발하였다(Gelber 1995).

이것은 전체 생존을 다양한 건강 상태로 나누고 건강상태의 평균 기간을 Kaplan-Meier 법을 이용해서 계산하고 각 상태의 가중치를 할당하고 치료군에서 가치를 계산한다. 이것은 인구집단의 관점을 채택하며 이득을 안녕상태로 계산하고 건강상태에 대한 선호도를 가중치를 통해서 고려한다.

Q-TWiST는 Quality-Adjusted Time Without Symptoms and Toxicity의 약자로 에이즈와 암의 무작위임상시험에서 치료방법들의 비교에 질보정 생존분석을 하는 것이다.

Q-TWiST 분석을 하려면 연구에서 치료와 질병의 특정 측면에 중점을 둔 삶의 질 차원에 대한 정의가 있어야 하고 질환의 간접비용과 정신적 상병에 대한 고려가 있어야 한다. 즉 임상적 결과지표, 부작용에 대한 정보, 삶의 질, 사회비용 그리고 선호도에 대한 정보를 얻어야 한다.

해당 자료와 주YG중치를 이용해서 각 치료에 대한 Q-TWiST 점수를 계산한다. 우선 각 시간간격에 대한 G중평가치를 계산하는데 이것은 각 치료간격에 대한 전반적 효용치와 같다.

locally advanced squamous cell carcinoma of the head and neck 환자에게 네 가지 서로 다른 방사선 분할 조사법의 효과를 quality adjusted survival (QAS)법으로 분석하였다(Konski 2009).

네 가지 분할조사법은 다음과 같다.

- 1) Standard Fractionation Radiotherapy (SFX)
- 2) Hyperfractionation Radiotherapy (HFX)
- 3) Accelerated Hyperfractionated Radiotherapy with Split (AHFXS)
- 4) Accelerated Fractionated Radiotherapy with Concomitant Boost (AFXC)

TWiST와 Q-TWiST는 아래와 같은 공식으로 계산하였다.

* TWiST = Survival - [Time in Toxicity (Tox)] - [Time in relapse (Rel)]

* Q-TWiST = TWiST + [(weight for Tox) × (Tox)] + [(weight for Rel) × (Rel)]

- Time in toxicity(독성기간)는 각각 다음과 같은 상태일 때로 정의하였으며 처음 해당 기준을 초과한 때로부터 기준이하로 낮아질 때 까지 시간을 조사하였다.

- 1) skin (≥ grade 2)
- 2) mucosa (≥ grade 2)
- 3) salivary gland (≥ grade 2)
- 4) esophagus (≥ grade 2)
- 5) larynx (≥ grade 3)
- 6) upper gastrointestinal (≥ grade 2)
- 7) spinal cord (≥ grade 2)
- 8) bone (≥ grade 3)
- 9) joint (≥ grade 3). Time

- 가중치는 환자 선호도로 계산하였다. 만일 재발에 대한 선호도가 동일하다고 했을 때 환자가 독성에 대해 높은 선호도를 보이는 경우가 그렇지 않은 경우보다 QAS 기간이 길어진다. 효용성(Utilities)은 해당 건강상태에 대해 전혀 선호하지 않을 때를 0(사망과 동일), 가장 높은 선호도를 보일 때 1(최선의 건강과 동일)로 표시하며 치료간 차이를 보는 기준을 독성 0.8, 재발을 0.5로 하였으며 통계적 유의성을 보이는 역치 분석을 하였다.

- HFX가 SFX에 비해 toxicity utility ≥ 0.57 relapse utility ≤ 0.72에서 QAS가 통계적으로 의미 있게 증가하였다. 다른 두 가지 방법은 SFX에 비해 통계적인 유의성이 없었다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

- 이 방법의 경우 가중치를 적용하는 건강상태에 대한 요소가 명백하지 않다는 점 그리고 으며 의사의 참조점(reference point)에 따라서 걸러지는 점 등이 문제점으로 지적되고 있다(Felli 2009).

2) A QALY-Based Approach to Benefit-Risk Modeling (FDA) : Incremental Net Health Benefit (INHB)

신약과 기존 약제의 INHB는 아래와 같이 표현된다.

$$\text{INHB} = (E2 - E1) - (R2 - R1)$$

E1, E2 : 각각 기존약제와 신약의 효과(QALY로 표현)

R1, R2 : 각각 기존약제와 신약의 위험(QALY로 표현)

만일 $(E2 - E1) > (R2 - R1)$ 이면 이득-위험 균형에서 이득 쪽이 높다고 할 수 있다. 이 모형의 경우 여러 가지 변이의 가능성이 있기 때문에 위험도에 대한 선호도로 보정할 필요성이 있으며 민감도 분석을 통해서 불확실성을 해소하여야 한다.

이 분석은 허가 전에 실시하며 허가 후에도 계속 이슈가 되면 허가 후에 실시할 수도 있다. 만일 안전성 정보를 더 모으기 위해 허가를 연기하면 모델일 통해 잠재적 이득 손실을 명백히 계산할 수 있다. 모든 방법론을 모든 제품에 대해 동일한 깊이로 시행하는 것은 아니다.

이 방법의 경우 위험도 추정에 몇 가지 문제를 초래할 수 있다. 이미 알려진 이상반응을 QALY로 표현하면 현재 임상 시험에서 거의 정보가 없을 가능성이 많으며 경미하지만 매우 중요할 가능성도 있다. 따라서 이에 대해서 다른 방법론이 도입될 가능성이 있다. 또한 알려지지 않은 심각한 이상 반응의 경우 허가 전에는 거의 문제가 되지 않기 때문에 이에 대한 보완이 필요할 수 있다.

이 방법의 장점은 매우 다양한 질병군에서 이미 많이 적용한 경험이 있다는 것이고 모든 단계에서 시행할 수 있다. 하지만 이 방법 자체를 잘 이해하지 못하거나 과학적으로 타당하다는 확신이 없으며 이상 사건보다는 이득에 더 많이 적용되었으며 QALY 자체는 위험에 대해 중립적이어서 위험에 대한 선호도의 차이를 반영하지 못한다.

3.2.3.3 "Principle of three" and the TURBO methods

The Council for International Organizations of Medical Sciences (CIOMS)⁹⁾는 "Principle of Threes"라는 방식으로 위험, 이득을 양적으로 표현하였

다.

이득, 위험비는 다음과 같은 형태로 표현한다. 'benefit score(cure rate × seriousness × duration of the disease) / individual AE scores(the incidence of the AE × the seriousness × the duration).

$$\frac{\text{이득점수(치유율} \times \text{심각도} \times \text{질병기간})}{\text{개별 이상반응(이상반응의 발생률} \times \text{심각도} \times \text{기간})}$$

이 방법은 질병의 적응증, 치료에 의한 개선, 치료 관련 이상 효과에 대한 질병 심각도, 기간, 발생률로 구분한다.

The European Agency for the Evaluation of Medicinal Products에서는 이 방법을 여러 가지 제한점을 가지고 있다고 평가하였는데 대체로 이득과 위험 목록이 포괄적이지 않고 일부만 통합되어 있으며 점수화의 기준이 애매하다는 것이다(CHMP 2007). 한마디로 매우 단순하지만 너무 인위적이라는 것이다.

3.2.3.4 Multi Criteria Decision Analysis (MCDA)

Multi-criteria decision analysis는 Mussen 등이 새롭게 제안한 위험-이득 분석 방법이다(Mussen 2007). 이 방법은 다양한 대상에 대한 선택을 할 때 생길 수 있는 불확실성을 해결하는데 초점을 맞추고 있다.

중요도에서 여러 차원을 가지는 다양한 선택권 사이의 저울질을 위해서 Multiple-attribute model이 사용된다. 이 방법은 산성비 정책이나 성과 연구 등 다양한 영역에서 유용성이 입증되었다.

MCDA는 이득과 위험으로 표현된 치료 특성에 대한 자료와 배열의 점수와 가중치를 투명하고 일관성 있는 방법으로 나열하고 이를 점수화 한다.

MCDA는 단순히 하나의 점수만을 제시하지는 않는다. 전체적인 하부 점수를 제시하며 불확실성에 따른 다양한 입력에 따라서 그 결과가 변화한다. 여러 MCDA 소프트웨어는 시각적인 혹은 점수화를 통해서 입력 혹은 하부 점수에 따른 변화를 볼 수 있게 해준다.

MCDA 전략은 의사 결정자로 하여금 자세한 조사가 필요한 영역과 집중할 부분에 대한 정보를 제공하며 점수 그 자체를 이용하는데 초점을 맞추지 않는다. 이런 전략은 여러 이해 당사자간의 이해와 차이를 해명하는데 이용되기도 하며 이를 통해서 투명한 의사 결정을 돕기도 한다.

9) Working CIOMS Group IV. Benefit-Risk Balance for Marketed Drugs: Evaluating Safety Signals. Geneva: Council for International Organizations of Medical Sciences; 1998.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Mussen에 의해 제안된 MCDA 과정은 다음과 같다.

1. 결정 상황(decision context)을 확립한다.
 - 대체로 이득-위험 분석에서는 적응증, 특정 적응증에서의 용량, 금기증의 세 가지 요소가 중요하다. 이러한 정보는 모두 의약품의 라벨링에 있다.
2. 평가할 옵션을 확인한다.
 - 평가할 약제는 무엇인지 그리고 무엇과 비교하는지에 대해서 확인한다.
3. 평가 목적과 기준을 확인한다.
 - 3.1. 개별 선택에 따른 결과를 평가하는 기준을 확인한다.
이득과 위험에 따른 기준은 다음과 같다.

이득

- 개별 주요 시험마다 효능과 임상적 적절성
- 중요 개별 시험의 이자인, 실행, 통계적 적절성
- 중요 개별 시험의 일차 종결점에 대한 임상적 적절성
- 중요 개별 시험의 목표 인구 대표성
- 중요 개별 시험의 효능 결과에 대한 통계학적 의미성
- 중요 개별 시험의 비일차적 결과에 대한 효능
- 비중요 개별 시험에 대한 효능
- 실제 임상에서 예상되는 환자의 순응도
- 중요 임상시험 결과들의 일관성

위험

- 이상 효과의 전체 발생률
- 심각한 이상 효과의 전체 발생률
- 이상 효과로 인한 중단 비율
- 특정 이상 효과의 발생률, 심각도, 기간과 가역성
- 하부 집단에서 안전성
- 다른 약제와 상호 작용
- off-label 사용으로 인한 위험의 잠재성
- 임상시험의 기간, 환자 범위에 따라 임상시험에서 드러나지 않는 잠재적 추가적 위험 정도
- 전임상연구에서 관찰된 드러나지 않는 잠재적 추가적 위험 정도
- 유사한 약제의 시험에서 드러난 드러나지 않는 잠재적 추가적 위험 정도

3.2. 개별 기준을 높은 수준 혹은 낮은 수준으로 묶어서 조직화함.

이 경우 가치수(value-tree)를 이용함(그림 16)

4. 기준에 대한 각 선택의 예상되는 수행도 평가한다.

4.1. 각 선택의 결과를 기술한다.

4.2. 결과에 대한 선택의 점수를 매긴다.

- 임상시험과 연장 연구, 안전 관련 데이터베이스의 자료 확인
- 각 기준의 척도 구성

4.3. 개별 결과에서 점수들의 일관성을 검토한다.

5. 개별 선택에서 각 결정의 상대적 중요도에 대한 가중치를 할당한다.

- 개별 기준에 대한 가중치를 배정함
- 주요 임상 시험과 다른 이득 기준에 기초해서 기저 수준 이득의 상대 가중치를 배정함
- 주요 시험과 다른 이득 기준의 상대 가치를 배정(bottom-up)
- 시험의 신뢰 기준을 고려함
- 위험 기준에 대한 상대 가중치 부여
- 이득과 위험에 대한 상대 가치 부여 (top-down)

6. 각 수준에 따른 가중치 점수를 계산하고 전체 가중치 점수를 계산한다.

- 모든 수준에서 기준의 가중치를 정규화 (모든 가중치를 100점 만점으로 환산)
- 각 옵션에서 정규화 가중치로 각 기준의 점수를 곱함
- 각 옵션에서 가중치 점수의 합인 전체 선호도 점수를 계산

7. 결과를 평가하고 민감도 분석을 한다.

- 가중치와 점수의 변이가 결과에 미치는 영향을 기술한다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

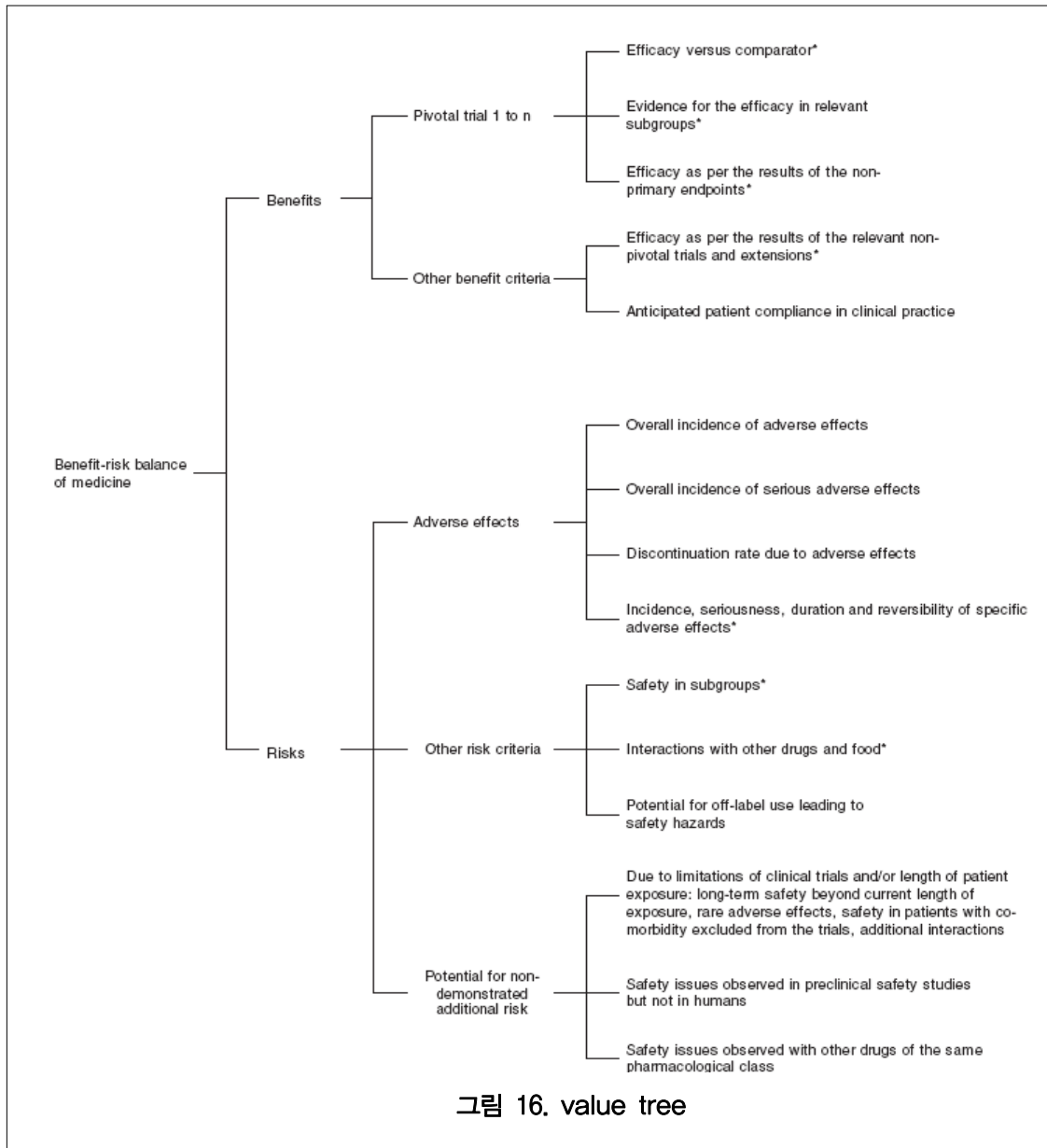


그림 16. value tree

3.2.3.5 방법간 비교 (5점 만점으로)

방법	범용성	방법론적 사용 용이	자료 획득의 용이성	이해 가능성	총점
LHH	2	4	3	4	13
RV-NNT	2	4	3	4	13
NNTT	2	2	3	3	10
Q-TWiST	4	3	3	3	13
INHB	5	4	3	4	16
TURBO	2	2	3	2	9
MCDA	5	3	3	4	15

- * 범용성 : 기구 등에서 논의되고 있거나 사용 중, 여러 그룹에서 이용, 여러 사례에서 분석됨 등
- * 방법론적 사용 용이 : 단계가 얼마나 복잡한지, 각 단계에서 복잡한 과정을 거치지 않는지
- * 자료 획득 용이성 : 계산에 필요한 자료를 쉽게 얻을 수 있는지 여부
- * 이해 가능성 : 해당 결과를 사용할 것으로 예상되는 집단에서 해당 내용을 얼마나 잘 이해할 수 있을지 여부

이상의 결과로 MCDA 혹은 INHB가 적절할 것으로 예상되며 NNT 관련 방법에 대한 실현 가능성을 모색할 필요가 있다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

3.3 급성 관상동맥 증후군의 표준치료법에 대한 NNT mapping

연구진이 NNT mapping으로 선정한 Acute Coronary Syndrome(이하 ACS)의 SIGN 가이드라인¹⁾에는 초기 질병 관리(initial management), ST 상승 급성 관상동맥 증후군에 대한 재관류(Reperfusion therapy for ST elevation acute coronary syndrome), 비침습적 검사와 위험 분류(Risk stratification and non-invasive testing), 침습적 검사와 혈관성형술(Invasive investigation and revascularisation), 조기 약제 치료(Early pharmacological intervention), 저산소증과 심장 쇼크의 치료(Treatment of hypoxia and cardiogenic shock), 환자 지지와 정보 충족(Patient support and information needs) 등 총 7개의 치료법이 제시되어 있다. SIGN 가이드라인은 각 치료법에 대해 대부분의 연구에서 근거등급과 함께 ARR 값과 RRR 값을 제시하였다. 연구진은 SIGN 가이드라인에서 근거등급이 1+ (잘 수행된 무작위임상시험들의 메타분석, 체계적 문헌고찰 혹은 비뚤림의 위험도(risk of bias)가 낮은 무작위임상시험) 이상인 치료법을 수용성 높은 치료법으로 보고 해당연구의 NNT 값에 대해 mapping을 하였다.

본 연구에서는 이중 초기 질병 관리와 조기 약제 치료 두 가지 치료법에 대한 NNT mapping을 제시하였다. 각 값은 부록에 제시하였다.

3.3.1 초기 약제 치료(Early pharmacological intervention)

초기 약제 치료는 ACS 발병 첫 12시간 이후 병원에서 퇴원하기까지의 약물치료를 의미한다. 퇴원 이후 장기간 치료에 대한 것은 다루지 않았다.

[그림 17]은 초기 약제 치료에 해당하는 전체 치료에 대한 NNT의 mapping이다. 경향성을 살펴보면 임상적 중요성(clinical importance) 또는 효용(utility)이라고 할 수 있는 preference의 값이 적을수록 NNT 값이 넓게 분포되어 있음을 볼 수 있다. 이는 임상적 중요성이 높은 치명적인 치료결과 일수록 NNT 값이 높은 치료법까지 수용가능하다는 것을 보여준다. 상대적으로 preference 값이 높은 치료결과에 대해서는 NNT 값이 낮은 치료법이 수용가능함을 보여준다.

Early Pharmacological intervention의 NNT mapping은 아래와 같다.

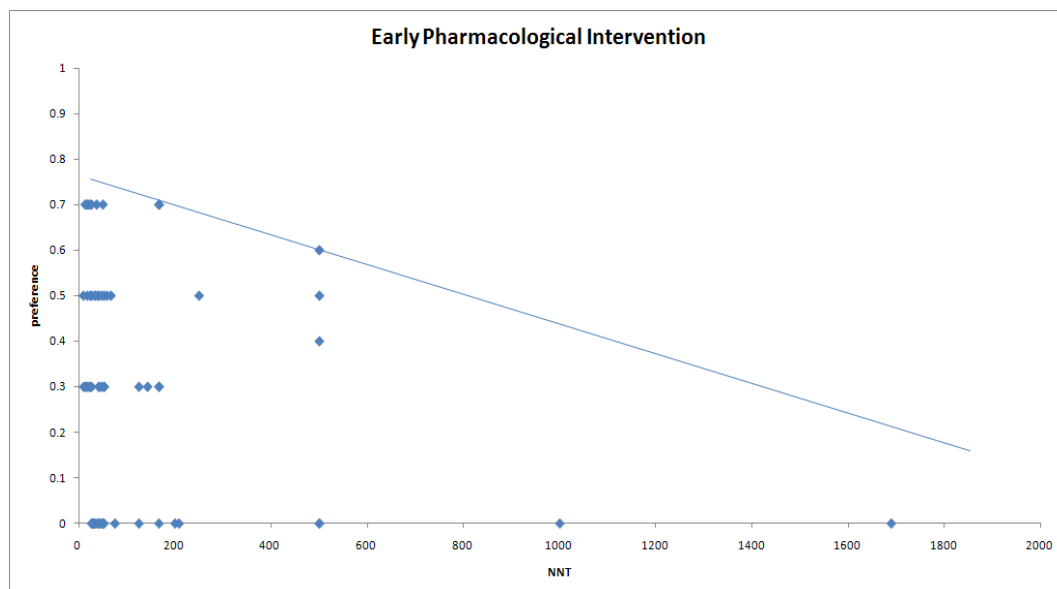


그림 17. ACS Early Pharmacological intervention NNT mapping

이를 경계선에 있는 자료를 이용하여 nonparametric 방법 중 하나인 lowess 방법으로 smoothing 하면 옵션의 변화에 따라 아래와 같은 곡선을 얻을 수 있다. (LOWESS (locally weighted scatterplot smoothing)는 산점도에서 두 변수간의 관계(relationship)를 smoothing하는 비모수적인 방법이다.) 하지만 이와 같이 modeling을 하기에는 사용되는 자료의 개수가 너무 적어서 어느 곡선의 형태로든 fitting 하기에는 무리가 있다.

이 곡선은 일종의 threshold로서 새로운 치료법을 평가할 때 이 곡선의 왼쪽에 mapping이 된다면 NNT와 치료결과의 측면에서 수용가능하다고 할 수 있고 오른쪽에 mapping 한다면 수용하기 어렵다고 가정할 수 있다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

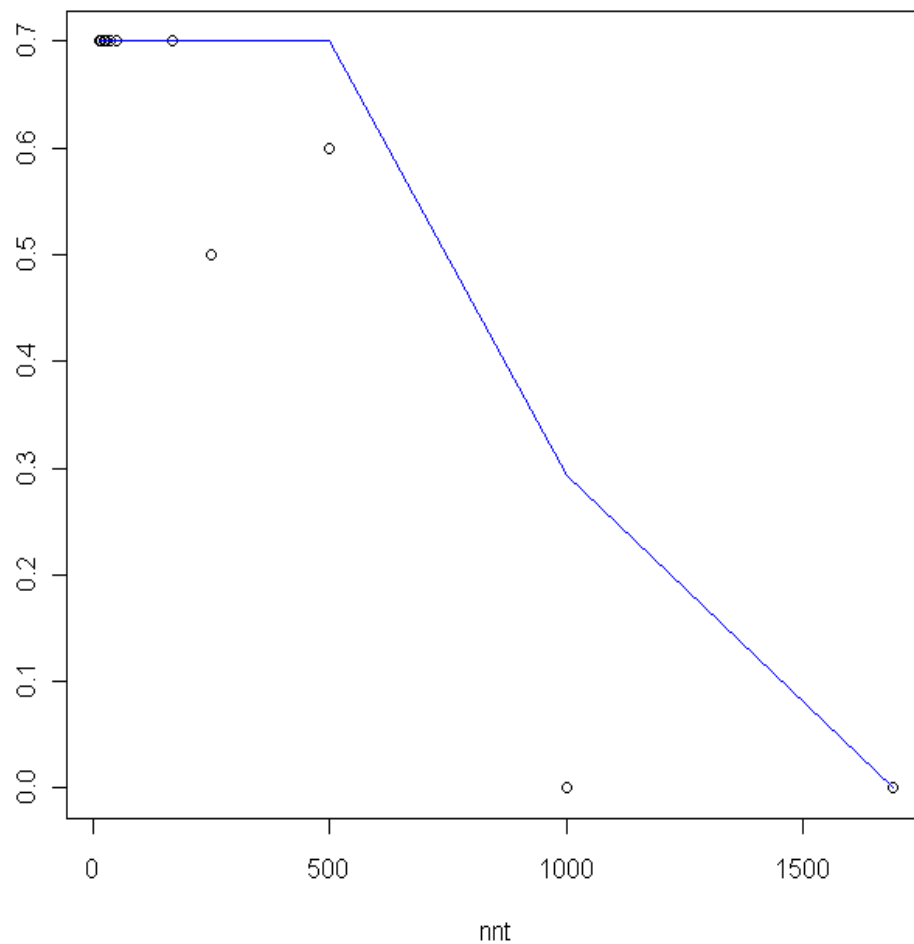


그림 18. Early Pharmacological intervention의 NNT mapping을 Lowess 방법으로 smoothing 한 결과

아래 각 질환 그룹별, 치료법별 mapping을 하였다. 아래의 mapping에서도 위와 비슷한 경향성을 보임을 알 수 있다.

Anticoagulant therapy의 NNT mapping은 아래와 같다.

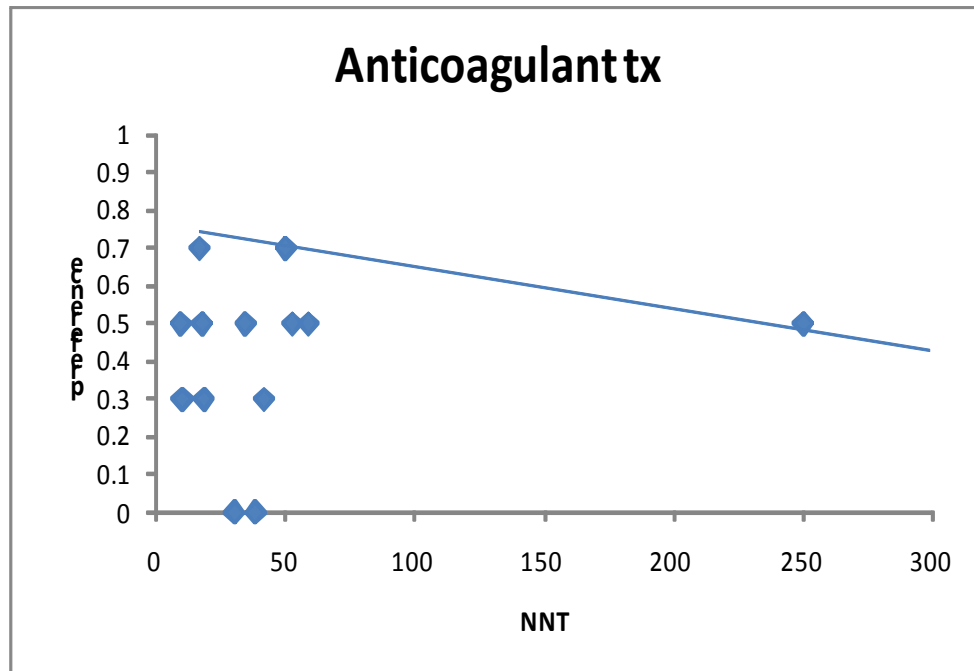


그림 19. ACS Anticoagulant therapy의 NNT mapping

Statin therapy의 NNT mapping은 아래와 같다.

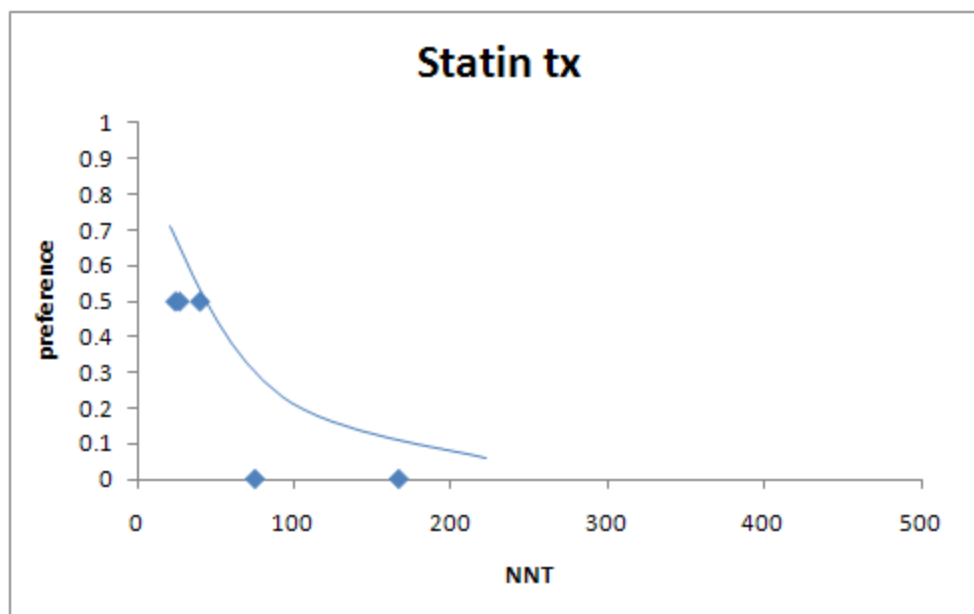


그림 20. ACS Statin therapy의 NNT mapping

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Beta Blocker therapy 와 Antianginal therapy의 NNT mapping은 아래와 같다.

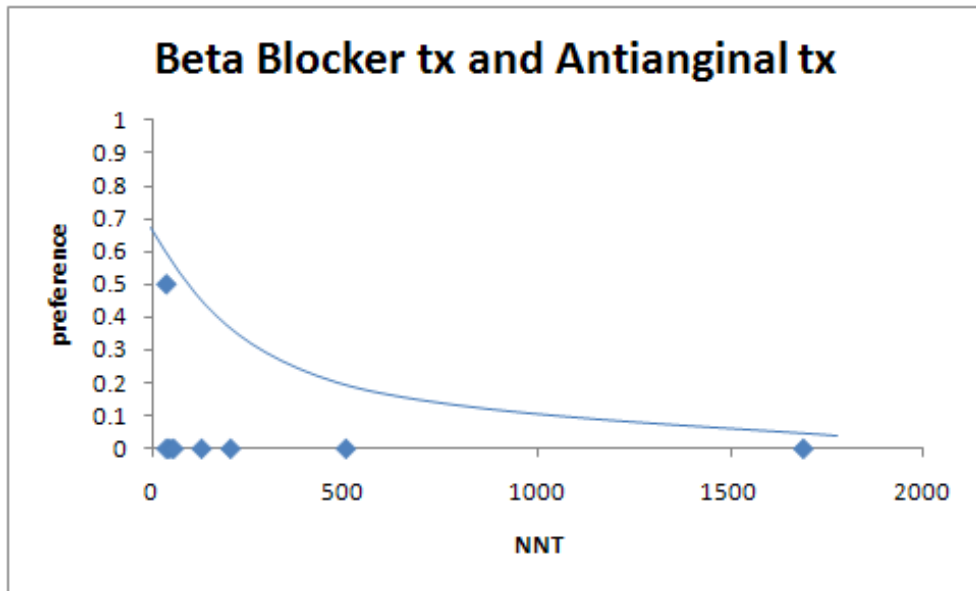


그림 21. ACS Beta Blocker therapy & Antianginal therapy의 NNT mapping

Angiotensin Receptor Blockers therapy의 NNT mapping은 아래와 같다.

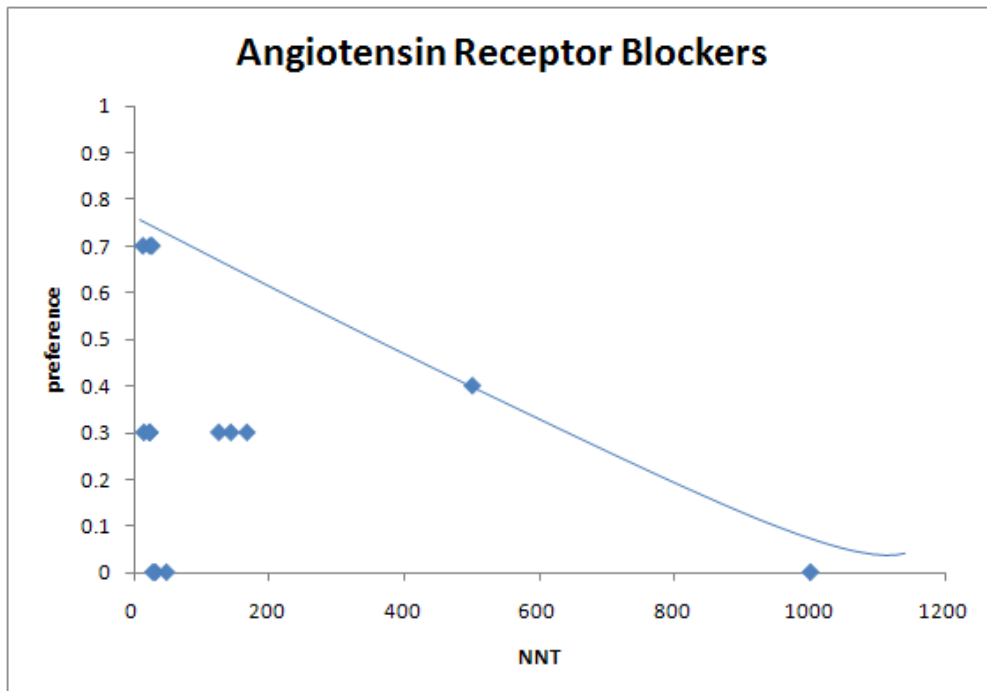


그림 22. ACS Angiotensin Receptor Blockers therapy NNT mapping

3.3.2 초기 질병 관리(Initial management)

초기 질병 관리는 ACS가 발병한 이후 첫 12시간 이내의 환자 관리하는 것을 의미한다. 초기 질병 관리의 NNT mapping 은 아래와 같다. (mapping 중 선은 연구자가 plotting 위에 임의로 선택하여 그린 것이다)

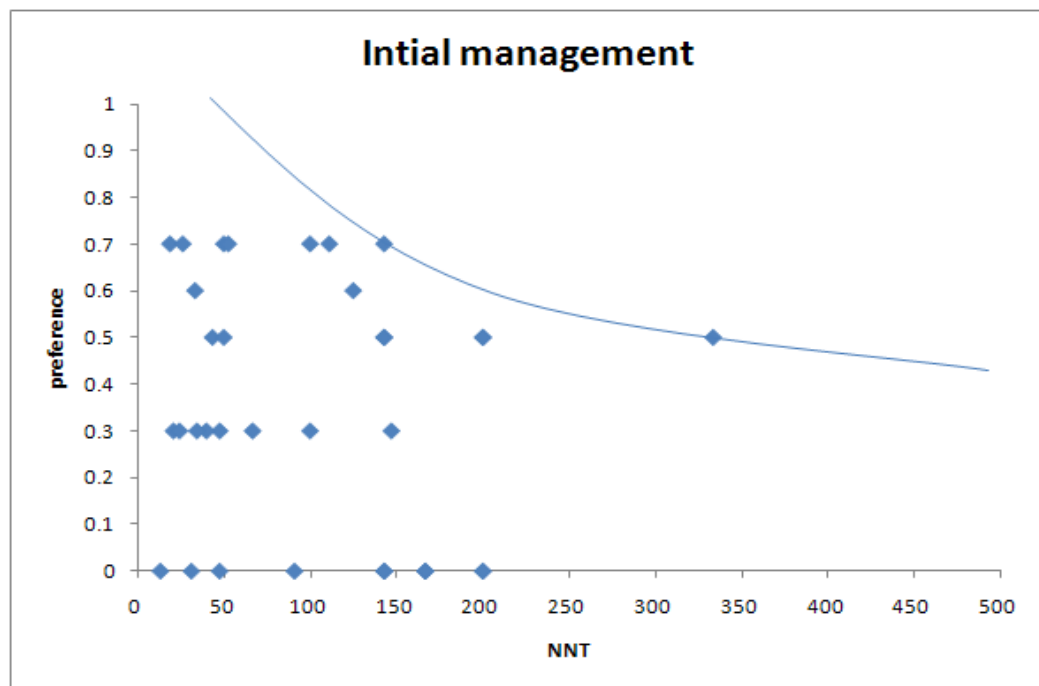


그림 23. ACS Initial management NNT mapping

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

초기 질병 관리 중 Non-ST elevation 의 경우는 아래와 같다.

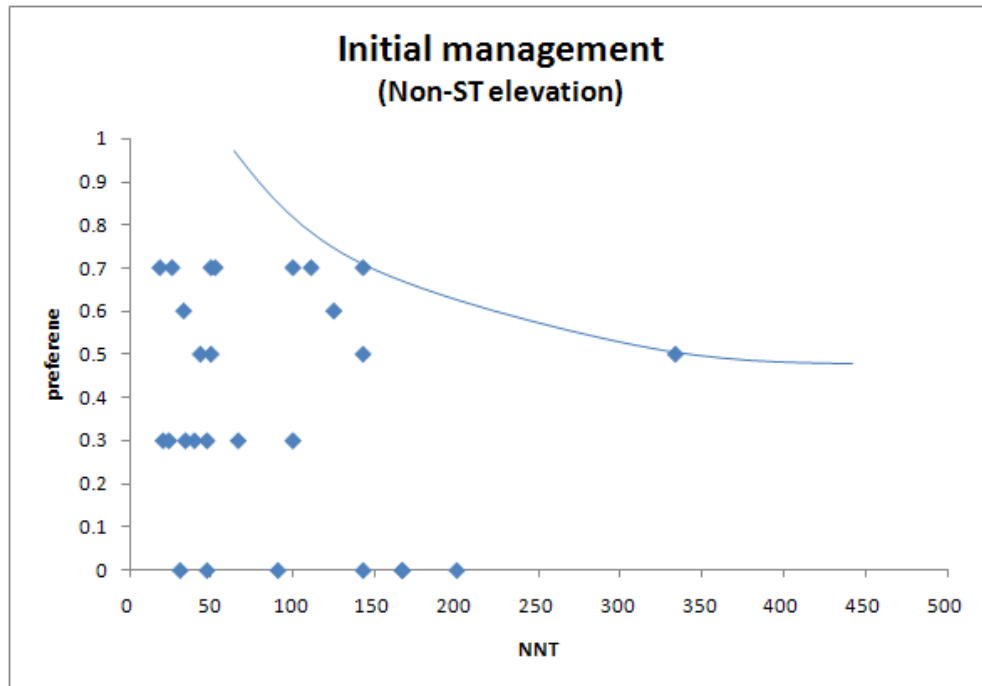


그림 24. Non-ST elevation ACS의 Initial management NNT mapping

초기 질병 관리 중 ST-elevation은 아래와 같다.

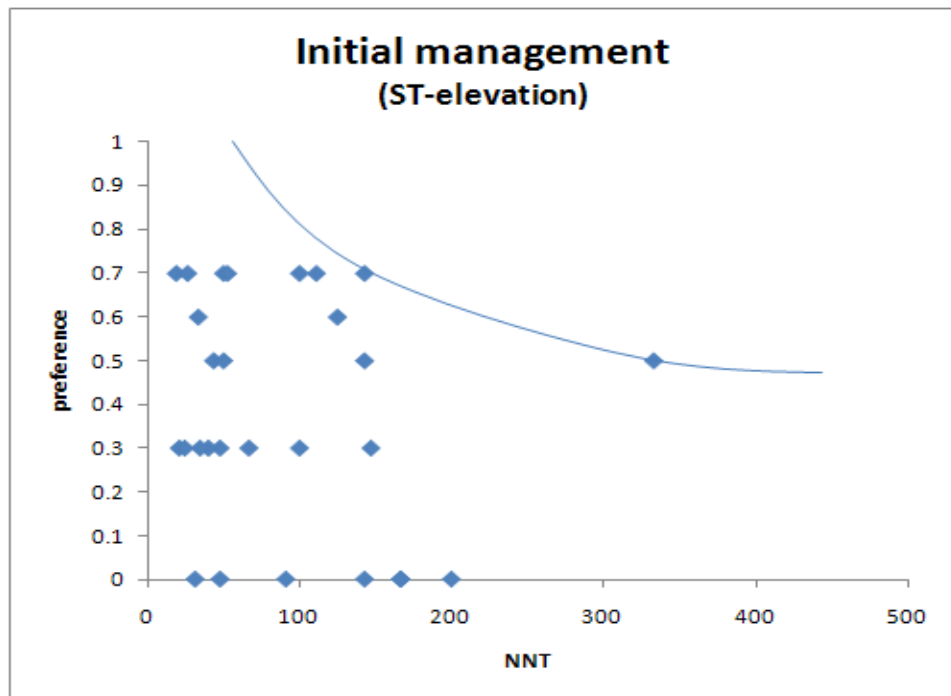


그림 25. ST elevation ACS의 Initial management NNT mapping

Antiplatelet therapy의 NNT mapping은 아래와 같다.

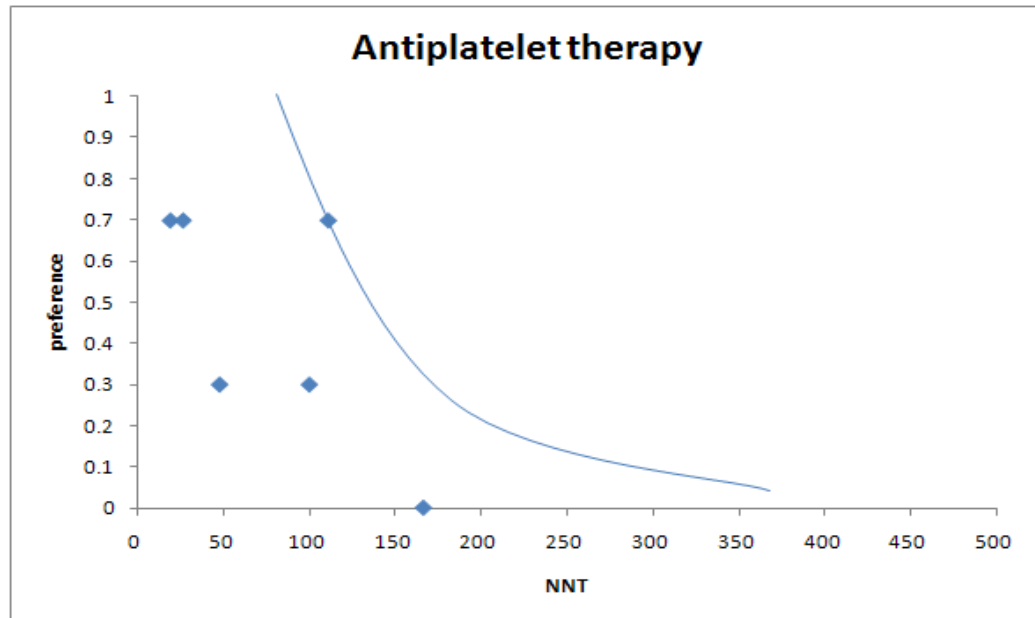


그림 26. ACS Antiplatelet therapy NNT mapping

Anticoagulant therapy의 NNT mapping은 아래와 같다.

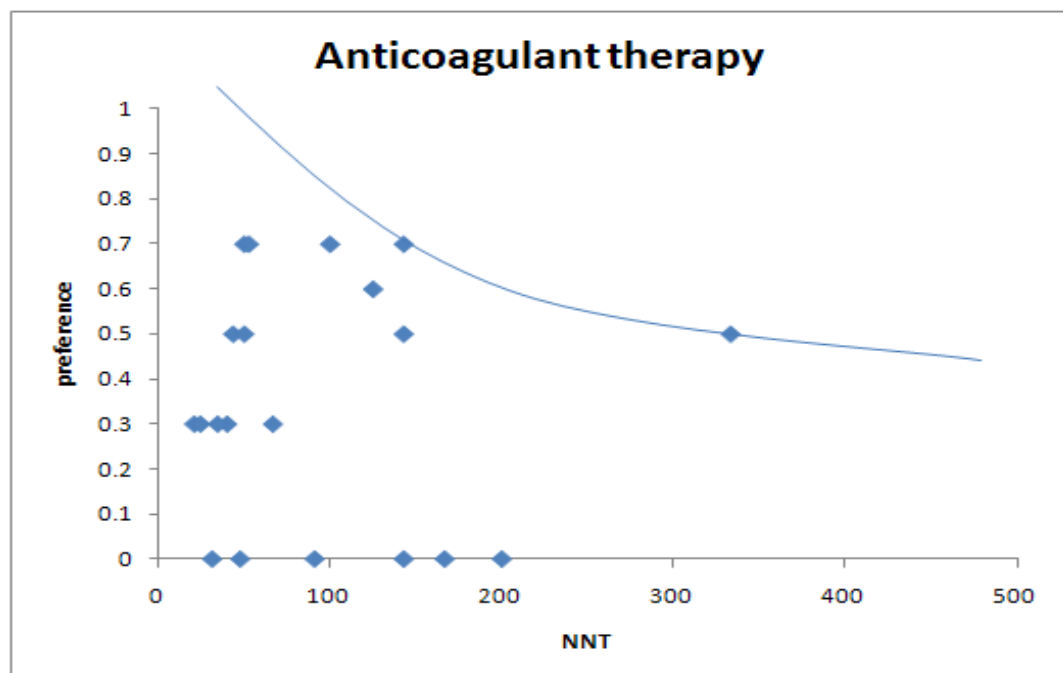


그림 27. ACS Anticoagulant therapy NNT mapping

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Beta Blockers의 NNT mapping은 아래와 같다.

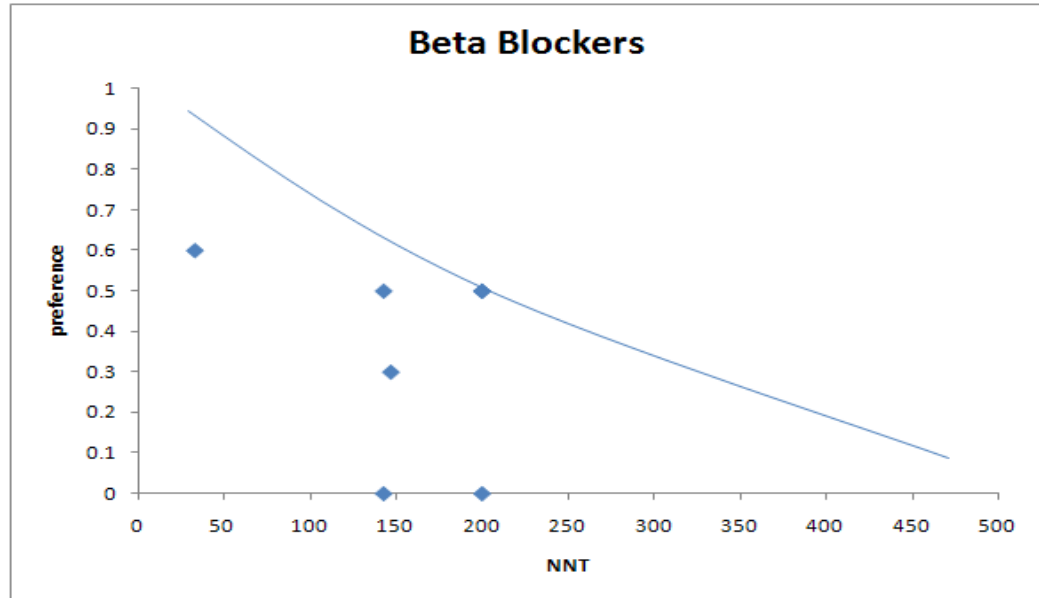


그림 28. ACS Beta Blocker therapy NNT mapping

3.3.3 ST 상승 급성관상동맥증후군에서 재관류요법(12시간 이내의 일차 PCI와 혈전용해에 초점을 둠)

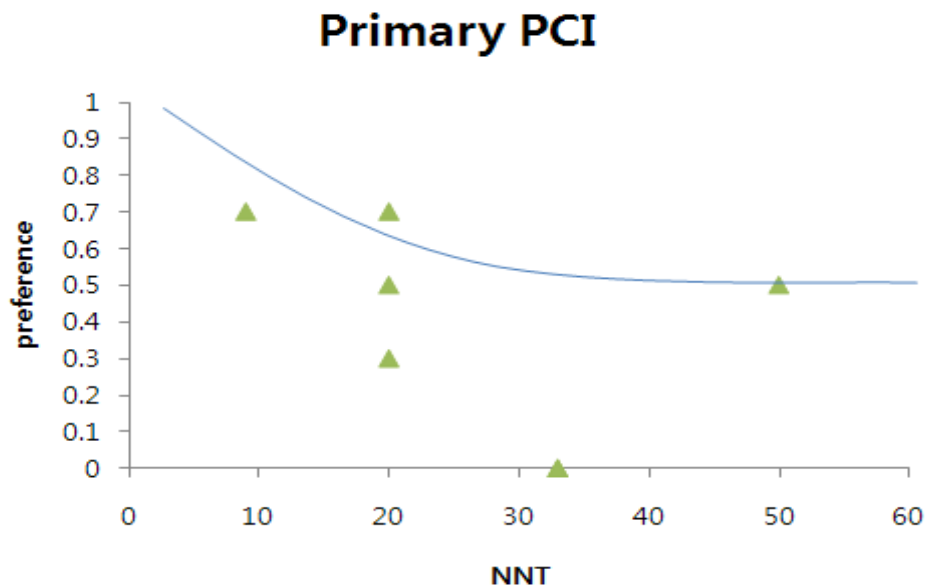


그림 29. ST 상승 ACS 환자에서 Thrombolysis에 비교된 Primary PCI의 NNT mapping

Adjuvant therapies for primary PCI

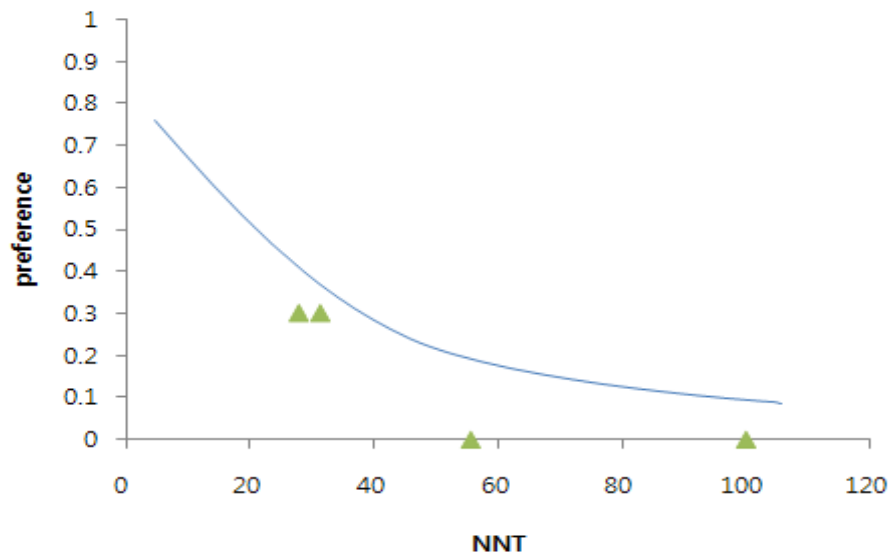


그림 30. Primary PCI로 치료받은 ST 상승 ACS환자에서 placebo와 비교된 Glycoprotein II b/III a receptor antagonists(abciximab)의 NNT mapping

3.3.4 침습적인 검사와 재관류요법

Routine coronary angiography and revascularization

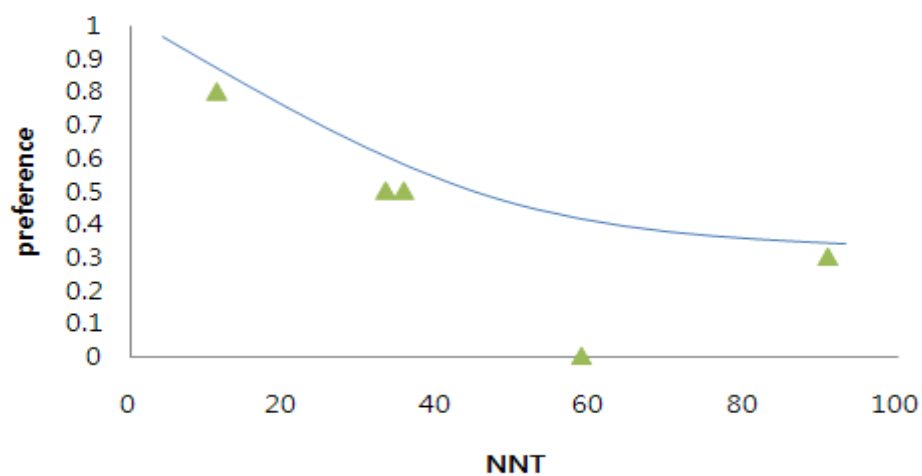


그림 31. non-ST 상승 ACS환자에서 conservative approach와 비교된 routine 관상동맥혈관성형술과 재관류요법

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

4. 고찰 및 결론

4.1 고찰

근거중심의사결정의 흐름을 보면 먼저 실험적이든 관찰적 연구이든 임상연구를 통한 근거의 생성과정(generation)이 있고 생성된 많은 일차 연구들에 대하여 체계적인 방식을 통하여 문헌들을 찾고 비평적 분석을 하며 가능한 경우 정량적 분석을 하여 근거를 합성(synthesis)하는 과정이 있게 된다. 이렇게 생성된 지식은 의사결정 주체를 통해 이 근거에 기초하여 여러 가치들을 고려한 가치판단(appraisal)의 과정을 거쳐 권고 혹은 결정을 하게 된다. 가치판단의 과정에서 과학적 정보나 지식에 대한 해석이 이루어지며 여기에 윤리, 사회 및 경제적 측면 등에 대한 고려를 하여 그 사회의 그 시점에서 현실을 반영한 최선의 의사결정을 하게 된다. 이러한 가치 판단의 과정에서 고려하는 사항들은 먼저 의학 적 근거로써 근거의 양과 질적 측면, 불확실성 그리고 이득과 위해의 종합적 판단이다. 또한 비용효과성, 이해당사자들의 견해, 수익자의 선호, 사회 윤리적 문제들 그리고 제공가능 현실성들에 대해 고려하게 된다.

가치판단의 가장 첫 번째 고려할 것이 의학적 근거이며 이 과학적 근거를 해석하는데 있어서 어떤 측면들이 고려해야 될지를 아는 것은 중요한 일이다. 먼저 치료의 효과를 측정하는 방법에 관해서는 짚의 특성에 따라 가지는 장점과 한계점들이 있다. 비교위험도(RR)이나 상대위험도감소(RRR)같은 지표들은 임상시험의 결과들을 요약하는데 유용한 지표들이며 이해하기 비교적 쉽고 매우 다른 임상적 상태에서도 RRR값은 같을 수 있으며 전향적 코호트 연구나 무작위대조군 임상시험에서 구할 수 있으며 그 외 단면 연구나 환자-대조군연구 같은 경우 사용할 수 없다. 결과의 확률에 기초한 것이므로 무작위 선정이나 대표성이 있는 표본이 선정된 경우에 한하여 사용할 수 있다는 제한점을 지니고 있다. 또한 현저한 치료의 효과와 절대수에 있어 매우 적은 효과를 분별하지 못할 수 있음으로 이러한 상대적 비교 특성을 지닌 결과 값만 제시하는 것은 위험할 수 있다. 따라서 이러한 점을 보완할 수 있도록 절대적 비교의 특성을 갖는 절대위험도감소나 NNT 같은 지표를 제시할 필요가 있고 특히 NNT는 한 바람직하지 않은 사건(事件; event)을 막기 위해 얼마의 노력을 들여야 할지를 보다 더 명확히 표현하며 동일비교 대상, 동일 위험군에서 유사한 시기의 결과 값을 측정할 경우 연구간, 치료법간 비교에 정보를 제공하게 된다. 그럼에도 불구하고 그러나 실제로 무작위배정 비교임상시험 결과들이 발표된 것을 분석한 결과 359편의 해당 연구들에서 8문헌에서 NNT를 18문헌에서 ARR의 보고를 함으로써 많은 경우 연구자들의 연구 결과 발표에 균형을 요한다(Nuovo 2002). 실제 우리나라에서도 NNT에 대한 인식이 아직 부족하고 그 개념을 이해하는 의사결정자들이 부족하므로 이 분야에

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

대한 교육과 지식의 확산이 절실하다고 생각한다.

물론 NNT가 지니고 있는 한계점도 또한 분명하여 기저 위험의 정도, 대조군의 종류, 임상연구에서 결과가 측정된 시기 등에 영향을 받게 되므로 치료간의 효과를 비교하는데 도움이 되지만 이러한 요소를 고려하지 않고 비교하는 오류를 범하지 말아야 한다. NNT 수치 자체는 다른 지표와 마찬가지로 그 결과값에 대한 가치가 반영되지 않은 수치이므로 NNT값만 갖고 판단할 수는 없으며 그 효용과 치료에 드는 비용 그리고 치료에 의한 부작용에 대한 고려를 해야 올바른 판단을 할 수 있다.

또한 각 지표가 갖는 특성 자체도 있지만 어떤 지표를 제시하는가에 따라 치료의 효과를 인식하는 정도도 달라질 수 있으므로 주의를 요한다. 많은 연구들에 있어서 정도의 차이가 있으며 상당히 이질적이긴 하나 일관되게 상대적 특성을 갖는 결과 측정법에 의해 의사결정하는 것이 절대적 특성을 지닌 결과 값 측정법에 의한 정보를 제시한 경우보다 효과에 대해 더 크게 인식된다는 연구에 대해 주의할 필요가 있다.

국가적으로 중요한 의사결정을 할 경우 그 의사 결정의 결과에 대해 후에 비평을 받게 되는 경우가 종종 생기게 되며 약제의 승인과정에 있어서 인허가의 의사결정에 대한 시비가 클 수 있다. 따라서 이런 의사결정을 하는 조직이나 위원회는 보다 객관적으로 치료법의 득과 실을 판단하는 방법이 필요하며 이에 대하여는 많은 연구들이 이루어져 왔으나 아직 어떤 방법이 최선의 방법인지 국제적인 합의된 부분은 부족하다. 하지만 득과 실에 대한 trade off는 의사결정이나 권고사항 결정과정에서 피할 수 없는 부분이며 다양한 이해 당사자들이 존재하는 보건의료분야의 특성상 투명하고 객관화 하시키는 노력을 기울이지 않을 수 없다. 이상적인 방법은 CHMP에서 제시한 바와 같이 가장 중요한 이득과 가장 심각한 위험을 쉽게 확인하게 해주는 것, 개별 이득과 위험에 대한 명백한 가중치를 부여하는 것, 근거의 강도와 확인된 불확실성에 대해 제시하는 것이며 이러한 원칙하에 각 치료법이 갖는 부차적인 이득과 위해를 종합적으로 반영하고 총합하여 결론을 내릴 수 있다면 좋겠지만 많은 변수를 고려할수록 방법론이 복잡해지고 실제 사용하기에 어려울 수도 있다.

미국 FDA의 새로운 이득-위험 분석의 조건으로 제시한 바에 따르면 약제 이득 위험 분석의 복잡성을 잘 포착하여야 하며 개별 환자에서 이득-비용의 특성을 양적으로 표현할 수 있어야 하고 새로운 정보가 나올 때 갱신할 수 있어야 한다. 또한 이득-위험 평가에 내제된 불확실성을 다루어야 하고 이득-위험 평가에서 비용의 문제를 해결할 수 있어야 한다. 양적 체계 도입의 비용 문제와 잠재적 이상 효과에 대하여 다루어야 하고 양적인 정보를 효과적으로 제시하고 교류할 수 있게 하여야 한다.

지금까지 연구를 토대로 향후 일상적으로 사용하기에 어렵지 않으면서 총괄적으로 이해득실을 평가할 수 있는 방법론의 도출이 필요하다.

4.2 결론

근거중심의사결정에 있어서 기존의 임상연구들에 대한 체계적이고 비평적인 방법을 통해 합성된 양질의 근거가 필요하며 이러한 연구들에 의하여 제시되는 결과의 측정 방법들은 각각 특징점과 제한점을 가지고 있다.

치료중재법의 치료효과 크기를 고려한 의사결정에서 효과의 크기를 비교 치료법에 대하여 상대적으로 표현하는 방법들과 차이를 절대적으로 표현하는 방식에 따라서 효과의 크기에 대한 인식이 달라질 수 있으므로 연구자는 결과값을 제시할 때 상대적 특성을 갖는 방법(RR, RRR)과 절대적 특성을 갖는 방법(ARR, NNT)값을 동시에 제시할 것을 권한다.

치료의 이득과 손실을 저울질하는데 있어 보다 객관화된 방법들이 도입되고 있다. 상황에 적절한 치료법에 대한 이득과 해로움을 측정하는 방법의 특성을 사용하여 의사결정과정에서 보다 객관화를 기할 필요가 있다.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

5. 참고문헌

박병주. 근거중심보건의료. 고려의학, 제 1판, 서울; 2009.

Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med. 2001 Apr 17;134(8):663-94.

Altman DG. Confidence intervals for the number needed to treat. BMJ 1998;317:1309-1312.

Bandolier. What is an NNT?, Implementing NNTs.

Barnett HJ, Taylor DW, Eliasziw M, et al. Benefit of carotid endarterectomy in patients with symptomatic moderate or severe stenosis. N Engl J Med. 1998;339:1415-1425.

Bland JM, Altman DG. The odds ratio. BMJ 2000;320:1468.

Carling CL, Kristoffersen DT, Montori VM, Herrin J, Schünemann HJ, Treweek S, Akl EA, Oxman AD. The effect of alternative summary statistics for communicating risk reduction on decisions about taking statins: a randomized trial. PLoS Med. 2009 Aug;6(8):e1000134. Epub 2009 Aug 25.

Cates CJ. Simpson's paradox and calculation of number needed to treat from meta-analysis. BMC Med Res Methodol. 2002;2:1. Epub 2002 Jan 25.

Committee for Medicinal Products for Human Use(CHMP). Reflection paper on benefit-risk assessment methods in the context of the evaluation of marketing authorisation applications of medicinal products for human use. Report of the CHMP working group on benefit-risk assess-

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

ment models and methods, London, 19 January 2007. cited 2010 Jan 11. Available from: <http://www.ema.europa.eu/pdfs/human/brmethods/1540407en.pdf>.

Committee for medicinal products for human use(CHMP). Reflection paper on benefit-risk assessment methods in the context of the evaluation of marketing authorisation applications of medicinal products for human use. Report of the CHMP working group on benefit-risk assessment methods. EMEA 2008.

Covey J. A meta-analysis of the effects of presenting treatment benefits in different formats. *Med Decis Making*. 2007 Sep-Oct;27(5):638-54.

Deeks JJ. When can odds ratios mislead? *BMJ* 1998;317:1155-6.

Fahey T, Griffiths S, Peters TJ. Evidence based purchasing: understanding results of clinical trials and systematic reviews. *BMJ*. 1995 Oct 21;311(7012):1056-9.

Felli JC, Noel RA, Cavazzoni PA. A multiattribute model for evaluating the benefit-risk profiles of treatment alternatives. *Med Decis Making*. 2009 ;29(1):104-15.

Frick MH, Elo O, Haapa K, Heinonen OP, Heinsalmi P, Helo P, Huttunen JK, Kaitaniemi P, Koskinen P, Manninen V, et al. Helsinki Heart Study: primary-prevention trial with gemfibrozil in middle-aged men with dyslipidemia. Safety of treatment, changes in risk factors, and incidence of coronary heart disease. *N Engl J Med*. 1987 Nov 12;317(20):1237-45.

Froud R, Eldridge S, Lall R, Underwood M. Estimating the number needed to treat from continuous outcomes in randomised controlled trials:

methodological challenges and worked example using data from the UK Back Pain Exercise and Manipulation (BEAM) trial. BMC Med Res Methodol. 2009 Jun 11;9:35.

Gelber RD, Cole BF, Gelber S, Goldhirsch A. Comparing treatments using quality-adjusted survival: the Q-TWiST method. Am Stat. 1995;49:161--9.

Gray A, Kehlet H, Bonnet F, Rawal N. Predicting postoperative analgesia outcomes: NNT league tables or procedure-specific evidence? Br J Anaesth. 2005 Jun;94(6):710-4. Epub 2005 Apr 15.

Griffith JM, Lewis CL, Hawley S, Sheridan SL, Pignone MP. Randomized trial of presenting absolute v. relative risk reduction in the elicitation of patient values for heart disease prevention with conjoint analysis. Med Decis Making. 2009 Mar-Apr;29(2):167-74.

Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, Schünemann HJ; GRADE Working Group. Rating quality of evidence and strength of recommendations: Going from evidence to recommendations. BMJ 2008 336(7652):1049-51.

Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. JAMA. 1994 Jan 5;271(1):59-63.

Halvorsen PA, Selmer R, Kristiansen IS. Different ways to describe the benefits of risk-reducing treatments. A randomized trial. Ann Intern Med 2007;146:848-56.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. West Wessex, UK.; John Wiley & Sons Ltd.; 2008.

Holden WL. Benefit-risk analysis : a brief review and proposed quantitative approaches. Drug Saf. 2003;26(12):853-62.

Institute of medicine. Understanding the benefits and risks of pharmaceuticals. workshop summary. the National academies press. 2007.

Julian P.T. Higgins, Sally Green. Cochrane handbook for systematic reviews of interventions. 2008.

Konski AA, Winter K, Cole BF, Ang KK, Fu KK. Quality-adjusted survival analysis of Radiation Therapy Oncology Group (RTOG) 90-03: phase III randomized study comparing altered fractionation to standard fractionation radiotherapy for locally advanced head and neck squamous cell carcinoma. Head Neck. 2009;31(2):207-12.

Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. Biol Psychiatry. 2006;59:990-6.

Kumana CR, Cheung BM, Lauder IJ. Gauging the impact of statins using number needed to treat. JAMA. 1999 Nov 24;282(20):1899-901.

Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med 1988;318:1728-33.

Lipsey , Wilson. Practical Meta analysis. http://psych.unl.edu/psycrs/971/meta/effect_sizes.pdf (접속일자 2010. 1. 11)

- Marx A, Bucher HC. Numbers needed to treat derived from meta-analysis: a word of caution. *ACP J Club*. 2003 Mar-Apr;138(2):A11-2.
- McAlister FA, Straus SE, Guyatt GH, Haynes RB. Users' guides to the medical literature: XX. Integrating research evidence with the care of the individual patient. Evidence-Based Medicine Working Group. *JAMA*. 2000 ;283(21):2829-36.
- McAlister FA. The "number needed to treat" turns 20 —and continues to be used and misused. *CMAJ* 2008;179:549-553.
- Moher et al. Guides for reading and interpreting systematic reviews. *Arch Pediatr Adolesc Med*. 1998;152(9):915-20.
- Montreuil B, Bendavid Y, Brophy J. What is so odd about odds? *Can J Surg*. 2005 Oct;48(5):400-8.
- Moxey A, O'Connell D, McGettigan P, Henry D. Describing treatment effects to patients: how they are expressed makes a difference. *J Gen Intern Med*. 2003;18:948-59.
- Mussen F, Salek S, Walker S. A quantitative approach to benefit-risk assessment of medicines - part 1: the development of a new model using multi-criteria decision analysis. *Pharmacoepidemiol Drug Saf*. 2007;16 Suppl 1:S2-S15.
- Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Intern Med*. 1992 Dec 1;117(11):916-21.
- Nuovo J, Melnikow J, Chang D. Reporting number needed to treat and absolute risk reduction in randomized controlled trials. *JAMA*. 2002 Jun

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

5;287(21):2813-4.

Osiri M, Suarez-Almazor ME, Wells GA, Robinson V, Tugwell P. Number needed to treat (NNT): implication in rheumatology clinical practice. *Ann Rheum Dis*. 2003 Apr;62(4):316-21.

Oxford league table of analgesics in acute pain. 2007 League table of number needed to treat (NNT) for at least 50% pain relief. cited 2009 Dec 31. Available from: <http://www.medicine.ox.ac.uk/bandolier/booth/painpag//Acutrev/Analgesics/Leagtab.html>.

Peat J, Barton B, Elliott E. *Statistics workbook for Evidence-based health care*. Wiley-Blackwell. West Sussex, UK. 2008.

Report of the CHMP working group on benefit-risk assessment models and methods. 2007. Doc. Ref. EMEA/CHMP/15404/2007.

Ringe JD, Doherty JG. Absolute risk reduction in osteoporosis: assessing treatment efficacy by number needed to treat. *Rheumatol Int*. 2009 Dec 25. [Epub ahead of print].

Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Evidence based Medicine; How to practice and teach EBM*. Edinburgh; Churchill Livingstone, 1997.

Samsa GP, Matchar DB, Goldstein L, Bonito A, Duncan PW, Lipscomb J, Enarson C, Witter D, Venus P, Paul JE, Weinberger M. Utilities for major stroke: results from a survey of preferences among persons at increased risk for stroke. *Am Heart J*. 1998;136:703-13.

Schechtman E. Odds ratio, relative risk, absolute risk reduction, and the number needed to treat--which of these should we use? *Value Health*.

2002 Sep-Oct;5(5):431-6.

Scottish Intercollegiate Guidelines Network. Acute coronary syndromes. 2007. <http://www.sign.ac.uk/pdf/sign93.pdf> (접속일자: 2010. 1. 11).

Sinclair JC, Cook RJ, Guyatt GH, Pauker SG, Cook DJ. When should an effective treatment be used? Derivation of the threshold number needed to treat and the minimum event rate for treatment. *J Clin Epidemiol*. 2001 Mar;54(3):253-62.

Sinclair JC. Weighing risks and benefits in treating the individual patient. *Clin Perinatol*. 2003 Jun;30(2):251-68.

Smeeth L, Haines A, Ebrahim S. Numbers needed to treat derived from meta-analyses - sometimes informative, usually misleading. *BMJ*. 1999 Jun 5;318(7197):1548-51.

Smith RS. 임상진료지침과 근거중심의학. 가정의학회지 1997;18(8S):87-97.

Scottish Intercollegiate Guidelines Network. Acute coronary syndromes. 2007. <http://www.sign.ac.uk/pdf/sign93.pdf>(접속일자: 2010. 1. 11)

Staquet MJ, Hays RD, Fayers PM. Quality of life assessment in clinical trial. 1st, Oxford university press, Oxford; 1998.

van 't Hof AW, Liem AL, de Boer MJ, Hoorntje JC, Suryapranata H, Zijlstra F. A randomized comparison of intra-aortic balloon pumping after primary coronary angioplasty in high risk patients with acute myocardial infarction. *Eur Heart J*. 1999;20:659-65.

Walter SD, Sinclair JC. Uncertainty in the minimum event risk to justify treatment was evaluated. *J Clin Epidemiol*. 2009;62:816-24.

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Wen L, Badgett R, Cornell J. Number needed to treat: a descriptor for weighing therapeutic options. Am J Health Syst Pharm. 2005 Oct 1;62(19):2031-6.

Working CIOMS Group IV. Benefit-Risk Balance for Marketed Drugs: Evaluating Safety Signals. Geneva: Council for International Organizations of Medical Sciences; 1998.

6. 부록

부록 1. The 2007 Oxford league table of analgesic efficacy (Oxford 2007)

Analgesic	Number of patients in comparison	Percent with at least 50% pain relief	NNT	Lower confidence interval	Higher confidence interval
Etoricoxib 180/240	248	77	1.5	1.3	1.7
Etoricoxib 120	500	70	1.6	1.5	1.8
Valdecoxib 40	473	73	1.6	1.4	1.8
Dipyrone 1000	113	79	1.6	1.3	2.2
Ibuprofen 600/800	165	86	1.7	1.4	2.3
Valdecoxib 20	204	68	1.7	1.4	2.0
Ketorolac 20	69	57	1.8	1.4	2.5
Ketorolac 60 (intramuscular)	116	56	1.8	1.5	2.3
Diclofenac 100	545	69	1.8	1.6	2.1
Piroxicam 40	30	80	1.9	1.2	4.3
Celecoxib 400	298	52	2.1	1.8	2.5
Paracetamol 1000 + Codeine 60	197	57	2.2	1.7	2.9
Oxycodone IR 5 + Paracetamol 500	150	60	2.2	1.7	3.2
Bromfenac 25	370	51	2.2	1.9	2.6
Rofecoxib 50	675	54	2.3	2.0	2.6
Oxycodone IR 15	60	73	2.3	1.5	4.9
Aspirin 1200	279	61	2.4	1.9	3.2
Bromfenac 50	247	53	2.4	2.0	3.3
Dipyrone 500	288	73	2.4	1.9	3.2
Ibuprofen 400	5456	55	2.5	2.4	2.7
Bromfenac 100	95	62	2.6	1.8	4.9
Oxycodone IR 10 + Paracetamol 650	315	66	2.6	2.0	3.5
Diclofenac 25	502	53	2.6	2.2	3.3
Ketorolac 10	790	50	2.6	2.3	3.1
Paracetamol 650 + tramadol 75	679	43	2.6	2.3	3.0
Oxycodone IR 10+Paracetamol 1000	83	67	2.7	1.7	5.6
Naproxen 400/440	197	51	2.7	2.1	4.0
Piroxicam 20	280	63	2.7	2.1	3.8
Lumiracoxib 400	370	48	2.7	2.2	3.5
Naproxen 500/550	784	52	2.7	2.3	3.3
Diclofenac 50	1296	57	2.7	2.4	3.1
Ibuprofen 200	3248	48	2.7	2.5	2.9
Dextropropoxyphene 130	50	40	2.8	1.8	6.5
Paracetamol 650 + tramadol 112	201	60	2.8	2.1	4.4
Bromfenac 10	223	39	2.9	2.3	4.0
Pethidine 100 (intramuscular)	364	54	2.9	2.3	3.9
Tramadol 150	561	48	2.9	2.4	3.6
Morphine 10 (intramuscular)	946	50	2.9	2.6	3.6
Naproxen 200/220	202	45	3.4	2.4	5.8
Ketorolac 30 (intramuscular)	359	53	3.4	2.5	4.9
Paracetamol 500	561	61	3.5	2.2	13.3
Celecoxib 200	805	40	3.5	2.9	4.4
Paracetamol 1500	138	65	3.7	2.3	9.5
Ibuprofen 100	495	36	3.7	2.9	4.9
Oxycodone IR 5 + Paracetamol 1000	78	55	3.8	2.1	20.0
Paracetamol 1000	2759	46	3.8	3.4	4.4
Paracetamol 600/650 + Codeine 60	1123	42	4.2	3.4	5.3
Paracetamol 650 + Dextropropoxyphene (65 mg hydrochloride or 100 mg napsylate)	963	38	4.4	3.5	5.6
Aspirin 600/650	5061	38	4.4	4.0	4.9
Paracetamol 600/650	1886	38	4.6	3.9	5.5
Ibuprofen 50	316	32	4.7	3.3	8.0
Tramadol 100	882	30	4.8	3.8	6.1
Tramadol 75	563	32	5.3	3.9	8.2
Aspirin 650 + Codeine 60	598	25	5.3	4.1	7.4
Oxycodone IR 5 + Paracetamol 325	149	24	5.5	3.4	14.0
Ketorolac 10 (intramuscular)	142	48	5.7	3.0	53.0
Paracetamol 300 + Codeine 30	379	26	5.7	4.0	9.8
Bromfenac 5	138	20	7.1	3.9	28.0
Tramadol 50	770	19	8.3	6.0	13.0
Codeine 60	1305	15	16.7	11.0	48.0
Placebo	>10,000	18	N/A	N/A	N/A

부록 2. Acute Coronary Syndrome의 치료효과크기 자료추출 결과

1. Initial management

Antiplatelet therapy	P	I	C	O	ARR	RRR	n	NNT	Utility
Aspirin	unstable angina	aspirin 300mg	placebo	the rate of vascular events(cardiovascular death, non-fatal MI, non-fatal stroke)	5.3%	46.0%	15,828	19	0.7
	acute MI	aspirin 300mg	placebo	the rate of vascular events	3.8%	30.0%	19,288	26	0.7
Combination aspirin and clopidogrel therapy	NSTE-ACS	aspirin 300mg +Clopidogrel 300mg	aspirin	cardiovascular death, stroke, MI	2.1%	20.0%	12,562	48	0.3
	STE-ACS	aspirin 300mg +Clopidogrel 300mg	aspirin	the rate of death, reinfarction, stroke	0.9%	9.0%	45,852	111	0.7
				the rate of death	0.6%	7.0%	45,852	167	0
Glycoprotein IIB/IIIA receptor antagonists	NSTE-ACS	Glycoprotein IIB/IIIA receptor antagonists	placebo	the odds of death or MI at 30days	1.0%	9.0%	31,402	100	0.3
Anticoagulant therapy	P	I	C	O	ARR	RRR	n	NNT	Utility
Unfractionated heprin	NSTE-ACS	Unfractionated heprin(UFH)	placebo	the combined end point of death or MI	2.5%	33.0%	1,353	40	0.3
	STE-ACS	Unfractionated heprin(UFH)	aspirin	the rate of re-infarction	0.3%		33,968	333	0.5
		Unfractionated heprin(UFH)	aspirin	the rate of death	0.5%		33,686	200	0
Low Molecular weight heparin (enoxaparin)	NSTE-ACS	LowMolecularweightheparin (enoxaparin)	unfractionated heparin (UHF)	to prevent one MI	0.8%	17.0%	11,128	125	0.6
				to prevent one extra revascularisation procedure	2.0%	12.0%	11,128	50	0.7

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Anticoagulant therapy	P	I	C	O	ARR	RRR	n	NNT	Utility
	STE-ACS	Low Molecular weight heparin(enoxaparin)	unfractionated heparin	MI	2.3%	41.0%	6,069	43	0.5
				recurrent ischemia	2.0%	30.0%	6,069	50	0.5
				death or MI	2.9%	26.0%	6,069	34	0.3
				death, MI or recurrent ischemia	4.8%	28.0%	6,069	21	0.3
		with alteplase or tenecteplase	unfractionated heparin	increase major bleeding	1.0%	44.0%	1,639	100	0.7
		Low Molecular weight heparin(enoxaparin)		death of recurrent MI	2.1%	17.0%	20,506	48	0
Direct Thrombin inhibitors		Direct Thrombin inhibitors	unfractionated heparin	increase major bleeding at 30 days	0.7%	53.0%	20,506	143	0.7
				re-infarction at seven days	0.7%	20.0%	17,073	143	0.5
Synthetic pentasaccharides	NSTE-ACS	Synthetic pentasaccharide, fondaparinux	enoxaparin	reduced risk of major bleeding	1.9%	48.0%	20,078	53	0.7
				short(30day) term mortality	0.6%	17.0%	20,078	167	0
				long term(180day) mortality	0.7%	11.0%	20,078	143	0
	STE-ACS	fondaparinux (subcutaneous inj. 2.5mg/day)	placebo or UFH	death or recurrent MI at 30days	1.5%	14.0%	12,092	67	0.3
				death rate at all times(9, 30, 180days)	1.1%	13.0%	12,092	91	0
		fondaparinux (subcutaneous inj. 2.6mg/day)	placebo or UFH	death	3.2%	21.0%	2,666	31	0
				death or recurrent MI	4.1%	23.0%	2,666	24	0.3

other therapy	P	I	C	O	ARR	RRR	n	NNT	Utility
Beta Blockers	NSTE-ACS	Beta Blocker		rate of progression to MI	3.0%	13.0%	4,700	33	0.6
	STE-ACS	intravenous beta blocker therapy		early(7 day) benefit in cardiovascular mortality	0.7%	15.0%	16,027	147	0.3
		oral beta blockade		re-infarction	0.5%	18.0%	45,852	200	0.5
				arrhythmic death	0.5%	17.0%	45,852	200	0
				mortality	0.7%	13.0%	52,645	143	0
				re-infarction	0.5%	22.0%	52,645	200	0.5
				cardiac arrest	0.7%	15.0%	52,645	143	0.5
Glycaemic Control		intensive metabolic control using insulin and glucose infusion	placebo	marked mortality benefit at one year	7.5%	29.0%	620	13	0

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

2. Early pharmacological intervention

Antiplatelet tx	P	I	C	O	ARR	RRR	n	NNT	Utility
Aspirin	Co		placebo	long term secondary preventive benefit	2.7%	37%		37	0.7
Clopidogrel	Non-ST elevation		placebo	CV Death/MI/stroke	2.1%	18%	12,562	48	0.3
	Medical therapy only		placebo	CV Death/MI/stroke	1.9%	19.0%	7,985	53	0.3
	revascularization		placebo	CV Death/MI/stroke	2.4%	17.3%	4,577	42	0.3
Anticoagulant tx	P	I	C	O	ARR	RRR	n	NNT	Utility
Coronary Heart Disease		warfarin : High-Intensity Oral Anticoagulant (OA)	no-aspirin	mortality	3.3%	19.4%		30	0
				MI	5.6%	34.1%		18	0.5
				thromboembolic complications	5.9%	59.6%		17	0.7
				Stroke	1.7%	44.7%		59	0.5
				Death, MI or Stroke	9.8%	32.6%		10	0.3
		warfarin : Moderate-Intensity OA	no-aspirin	mortality	2.6%	10.7%		38	0
				MI	10.4%	42.8%		10	0.5
				Stroke	2.9%	50.9%		34	0.5
				Death, MI or Stroke	2.4%	7.1%		42	0.3
		High- or Moderate-Intensity OA	aspirin	major bleeding	-2.7%	-270.0%		(37)	0.7
				death, MI, stroke	5.4%	50.0%		19	0.3

Anticoagulant tx	P	I	C	O	ARR	RRR	n	NNT	Utility
		high- or moderate-intensity OA and Aspirin	aspirin	major bleeding	-1.6%	-94.1%		(63)	0.7
		warfarin+aspirin	aspirin	MI	1.9%	44.0%		53	0.5
				ischemic stroke	0.4%	54.0%		250	0.5
				coronary revascularisation	2.0%	20.0%		50	0.7
				major bleeding	-0.9%	150.0%		(111)	0.7

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Statin tx	P	I	C	O	ARR	RRR	n	NNT	Utility
	CHD		placebo	CHD mortality	1.3%	29%	30,817	75	0
	healthy population		placebo	major coronary events in the primary prevention		34%	13,200		0.6
	CHD		placebo	major coronary events in the secondary prevention		30%	17,617		0.7
	healthy population	(WOSCOP)	placebo	major coronary event	2.5%	29%	6,595	40	0.5
	healthy population	(WOSCOP)	placebo	coronary mortality	0.6%	33%	6,595	167	0
	healthy population	(AFCAPS/TexCAPS)	placebo	major coronary event	4.1%	37%	6,605	24	0.5
	healthy + CHD	(CTT META)	mainly placebo	5 year, major vascular event	3.7%	79%	90,056	27	0.5
Beta Blocker tx & Antianginal tx	P	I	C	O	ARR	RRR	n	NNT	Utility
BetaBlockertx	ACS without clinical MI		(control)	MI	3.0%	9.4%		33	0.5
	ACS with clinical MI			2 year death			24,974	42	0
	clinical MI with LVF	Carvedilol	placebo	all-cause mortality	3.0%	23%	1,959	33	0
Nitrates and calcium channel blockers	acute MI	Captopril	placebo	5wk mortality	0.5%	6.5%	58,050	200	0
	acute MI	Mononitrate	placebo	5wk mortality	0.2%	2.7%	58,050	500	0
	acute MI	Magnesium	placebo	5wk mortality	-0.4%	-5.5%	58,050	(250)	0
	acute MI	transdermal GTN	open control	6wk mortality	0.8%	11.3%	19,394	125	0
	MI	diltiazem(calcium channel blocker)	placebo	death	0.1%	0.0%	2,466	1689	0
				1st recurrent cardiac event	1.9%	0.0%	2,466	52	0

ACE Inhibitors	P	I	C	O	ARR	RRR	n	NNT	Utility
	ACS without clinical MI :highriskpatientswithvascular disease	ramipril	placebo	all cause mortality, MI, stroke	3.8%	21.3%	9,297	26	0.3
				death from cardiovascular cause	2.0%	24.7%	9,297	50	0
				MI	2.4%	19.5%	9,297	42	0.5
				Stroke	1.5%	30.6%	9,297	67	0.5
	stable CHD	Perindopril 8mg	placebo	cardiovascular death, MI, cardiac arrest	1.9%	19.2%	12,218	53	0.3
	stable CHD	trandolipril	placebo	Primary : death from cardiovascular causes, nonfatal MI, CABG, PCI	0.6%	2.7%	8,290	167	0.3
				Death from cardiovascular causes	0.2%	5.4%	8,290	500	0
				nonfatal MI	0.0%	0.0%	8,290		
				CABG	0.6%	8.5%	8,290	167	0.7
				PCI	-0.4%	-3.3%	8,290	(250)	0.7
	acute MI	ACE inhibitor	mainly placebo	death	0.5%	6.3%	98,483	208	0
				nonfatal cardiac failure	0.6%	3.9%	98,483	167	0.7
				ventricular fibrillation	0.2%	6.1%	98,483	500	0.6
	heart failure or left-ventricular dysfunction	ACE inhibitor	placebo	death	3.8%	14.2%	12,763	26	0
				re-infarction	2.1%	19.1%	12,763	48	0.5
				readmission for CHF	5.2%	27.5%	12,763	19	0.7
				death/MI/readmission for CHF	7.2%	17.6%	12,763	14	0.3
				stroke	0.2%	5.1%	12,763	500	0.5

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

Angiotensin Receptor Blockers	P	I	C	O	ARR	RRR	n	NNT	Utility
	acute MI	Losartan	Caprtopril	all-cause mortality	-1.8%	-11.0%	5,477	(56)	0
				MI	-0.1%	-0.7%	5,477	(1000)	0.5
				sudden cardiac death/resuscitated cardiac arrest	-1.3%	-17.6%	5,477	(77)	0
	MI	Valsartan	Caprtopril	Death from cardiovascular causes(A)	0.1%	0.6%	9,818	1000	0
				A or MI	0.7%	3.0%	9,818	143	0.3
				A or heart failure	0.2%	0.7%	9,818	500	0.4
				A, MI, heart failure	0.8%	2.5%	9,818	125	0.3
				A, MI, HF, resuscitation after cardiac arrest, stroke	0.6%	1.8%	9,818	167	0.3
	chronic heart failure	candesartan	placebo	Cardiovascular death or hospital admission for CHF	4.3%	12.5%	7,599	23	0.3
				cardiovascular death	2.1%	10.3%	7,599	48	0
				hospital admission for CHF	4.3%	17.8%	7,599	23	0.7
	CHF	candesartan	placebo	Cardiovascular death or hospital admission for CHF	7.0%	17.5%	2,028	14	0.3

Angiotensin Receptor Blockers	P	I	C	O	ARR	RRR	n	NNT	Utility
				cardiovascular death	3.2%	12.9%	2,028	31	0
				hospital admission for CHF	7.8%	27.7%	2,028	13	0.7
	CHF	candesartan	placebo	Cardiovascular death or hospital admission for CHF	4.4%	10.4%	2,548	23	0.3
				cardiovascular death	3.6%	13.2%	2,548	28	0
				hospital admission for CHF	3.8%	13.6%	2,548	26	0.7

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

3. Reperfusion therapy for ST elevation ACS (within the first 12 hours - focusing on both primary PCI

P	I	C	O	ARR(%)	RRR(%)	n	NNT	Utility	comment
ST elevation ACS	Primary PCI	thrombolysis	short term mortality	3.0%	36%	1,778	33	0	
			long term mortality	3.0%	38%	828	33	0	
			Stroke	2.0%	64%	1,778	50	0.5	
			Re-infarction	5.0%	59%	1,570	20	0.5	
			Recurrent ischaemia	11.0%	59.0%	1,156	9	0.7	
			Death or non-fatal re-infarction	5.0%	44.0%	1,494	20	0.3	
			Need for CABG	5.0%	36.0%	892	20	0.7	
	Adjuvant therapies for primary PCI								
ST elevation ACS treated with primary PCI	Glycoprotein II b/III receptor antagonists (abciximab)	placebo	30-day of death, reinfarction, or urgent or ischemia-driven TVR	3.6%	46.0%	3,266	28	0.3	A논문
		placebo	6 month of death, reinfarction, or urgent or ischemia-driven TVR	3.2%	20.0%	3,266	31	0.3	A논문
		placebo or none = control	30-day mortality	1.0%	29.0%	27,115	100	0	A논문
		placebo or none = control	6-12 month mortality	1.8%	29.0%	27,115	56	0	B논문
		placebo or none = control	30-day reinfarction	0.9%	38.1%	27,115	111	0.5	B논문
	Intracoronary stenting								
ST elevation ACS treated with primary PCI	Intracoronary stenting	Isolated balloon angioplasty	reinfarction at 12 months	1.2%	33.0%		83	0.5	
			Target vessel revascularization at 12 months	14.4%	52.0%		7	0.5	

P	I	C	O	ARR(%)	RRR(%)	n	NNT	Utility	comment
	Thrombolytic therapy								
STelevationACS	Thrombolytic therapy	placebo	35-day mortality	1.9%	18.0%		53	0	
	Thrombolytic therapy in the very early phase (≤ 2 hrs)	placebo (Tx > 2 hrs)	35-day mortality	24.0%			4	0	
	Service delivery								
acute MI	Transfer of pts to interventional centres	thrombolytics	death, re-infarction, stroke at 30 days	6.0%	40.0%		17	0.3	A
			primarily driven by a reduction in re-infarction	4.7%	75.0%		21	0.5	A
			death, re-infarction, stroke at 30 days	7.0%	45.0%		14	0.3	B
			primarily driven by a reduction in re-infarction	1.7%	55.0%		59	0.5	B
	Choice of thrombolytic agent	Choice of thrombolytic agent	Choice of thrombolytic agent			40,539			
acute MI	administration over 90 minutes : accelerated regimen of alteplase	streptokinase	mortality	1.1%	14%		91	0	
pts who fail to reperfuse after thrombolysis	Rescue PCI after failed thrombolysis	conservative therapy	early severe heart failure	8.0%	68%		13	0.5	
		conservative therapy	1-year mortality	5.0%	38%		20	0	
		delayed PCI (at mean of 12 days post-infarction)	death, re-infarction, revascularization, ischaemic events at 6 months	25.0%	49%		4	0.3	

치료법의 수용에 있어서 치료적 효과 크기에 대한 의학-사회적 의미

4. Invasive investigation and revascularization

P	I	C	O	ARR(%)	RRR(%)	n	NNT	Utility
Non-ST elevation ACS	routine coronary angiography and revascularization	conservative approach	death or MI after discharge	1.1%		9,212	91	0.3
			severe angina after discharge	2.8%	23%	9,212	36	0.5
			rehospitalization after discharge	8.8%	34%	9,212	11	0.8
			MI	3.0%	26%	2,457	33	0.5
			1 year mortality	1.7%	43.0%	2,457	59	0
moderate risk patients with Non-ST elevation ACS	early invasive investigation and revascularisation		death, MI, refractory angina at 4 months	4.9%	59.0%	1,810	20	0.3
ST elevation ACS	PCI or CABG	conservative (ischaemia-driven) treatment	death, MI or revascularization	12.0%	56.0%	500	8	0.3

부록 3. 근거중심 의사결정 방법론 심포지엄

- Benefit & Harm, Its trade off -

2010년 2월 8일(월요일) 14:00-18:00

Session 1

치료적 중재 효과 측정의 올바른 이해

이상무 연구위원(한국보건의료연구원 보건의료분석실)

절대적, 상대적 치료효과 크기 지표의 장단점 분석, 제시된 치료효과 크기에 따른 의사결정자들의 효과에 대한 인식이 달라지는 것에 대한 사례발표

Session 2

의약품 허가에서의 Risk-benefit 상대비교 국제 동향

이의경 교수(숙명여자대학교 임상약학대학원)

Risk-Benefit의 상대 비교에 대한 개념에 대해 알아보고, WHO CIOMS Working groups의 약제 안전성 평가를 위한 주요 고려 항목 및 방법 제시, 유럽의 EMEA와 미국 FDA의 Benefit-risk analysis 동향 제시

Session 3

근거중심 의사결정에서의 trade-off

김수영 교수(한림대학교 의과대학)

근거중심 의사결정에서 이득과 위해의 연구결과에 대한 권고 statement 기술에 대한 설명, 체계적 문헌고찰에서의 근거의 종합요소(근거의 질, 이득과 위해의 balancing, 가치와 선호도, 자원이용), GRADE 과정에 대한 설명

Session 4

약제의 Risk-Benefit 평가 방법

RTI international (F. Reed Johnson)

이득과 위해의 trade-off 선호도를 계량적으로 측정하는 방법 및 8개 질환에 대한 Benefit-risk preference 연구에 대한 설명

2010년 2월 10일(수요일) 13:00-17:00

Session 1

Quantifying Patient and Physician Benefit-Risk Tradeoff Preferences Using Conjoint Analysis

RTI international (F. Reed Johnson)

Stated-Preference methods(conjoint analysis)를 이용한 환자와 임상 의사의 이득과 위해 간의 trade off preference를 계량적으로 측정하는 연구 방법론 설명 및 실제 연구 수행에 있어서 survey design, survey administration, data analysis에 대한 설명

발행일 2009. 12.

발행인 허대석

발행처 한국보건의료연구원

이 책은 한국보건의료연구원에 소유권이 있습니다.
한국보건의료연구원의 승인없이 상업적인 목적으로
사용하거나 판매할 수 없습니다.