

# Kritisk evaluering af Patientsikkert Sygehus DK: Anmeldelse og metodiske refleksioner

Lars Ehlers

Professor, DCHI, AAU



# Agenda

1. Patientsikkert Sygehus – kort fortalt
2. Hvad sagde evalueringen? – kort fortalt
3. Kritiske refleksioner over valg af metode, herunder diskussion af
  1. Complex vs. non-complex interventions
  2. Evaluation vs. research
4. Kan der opstilles 'gode råd' til evaluering af fremtidige komplekse interventioner?

# CV - Lars Ehlers

1992 Cand. Oecon., Aarhus Universitet

1996 Ph.D. Aarhus Universitet

1996 Økonom, Tele Danmark Hovedkontor og  
Tele Danmark Net

1999 Sektionsleder, Tele Danmark Net

2001 Forældreorlov

2002 Konsulent og ekstern lektor, Handels-  
højskolen i Århus

2003 MTV-konsulent, Center for Folkesundhed,  
klinisk lektor Aarhus Universitet

2009 Professor, DCHI, Aalborg Universitet

# Social Science:1992-1996

- Thesis: Alternative organizational forms in hospitals
  - a comparative analysis of efficiency
  - focus on 'market-like institutional forms' as a complex intervention
- Case-study research method (Yin 1984)
  - "...the need for case studies arises out of the desire to understand complex social phenomena"
  - Theory and program theory
  - Multiple sources of information (quantitative and qualitative)



# Institutional research in the 90s

- ☞ No optimal way of organization
  - Many organizational forms may suffice
  - Some organizational forms are insufficient
  - It all depends (....on the situation)
- ☞ Reforming health care is extremely difficult.
  - Many reforms don't achieve the intended aims.
  - Reform typically requires the use of multiple instruments over a longer time period
- ☞ The outcome of a complex reform is difficult to predict and difficult to study

# Design depend on ...(question)

<b>Strategy</b>	<b>Form of Research Question</b>	<b>Requires Control of Behavioral Events?</b>	<b>Focuses on Contemporary Events?</b>
<b>Experiment</b>	how, why?	Yes	Yes
<b>Survey</b>	who, what, where, how many, how much?	No	Yes
<b>Archival analysis</b>	who, what, where, how many, how much?	No	Yes/No
<b>History</b>	how, why?	No	No
<b>Case study</b>	how, why?	No	Yes

**Figure 1.1** Relevant Situations for Different Research Strategies  
SOURCE: COSMOS Corporation.

# Solution depends on ... (context)

Markets

Customers

Finance and Accounting

Management & development of people

Ethics & responsible management

Organisational behaviour

Information systems & technologies

Business policy and strategy

Business environment

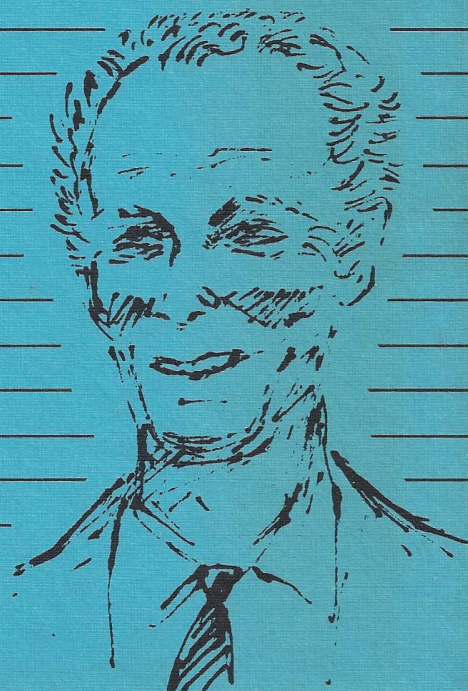
Knowledge management

Contemporary & pervasive issues



# DET LIGGER STADIG I LEDELSEN

Essays til  
Vagn Madsen



I anledning af dr.  
oecon Vagn  
Madsens fratræden  
i 1983 efter 30 år  
som professor ved  
Afd. for  
Virksomhedsledelse,  
Aarhus Universitet



# Health Science: 2003 - 2009

## HTA:

- Hierarchy of evidence



# Kun ca. 10-20% af sundhedsfaglige interventioner er evidensbaserede (Sackett)

## Det hvide snit

- En af de største fejltagelser i lægevidenskabens historie
- Ca. 150.000 operationer



Figure 141 (a) Case 262. A six year old girl whose destructive behavior was almost intolerable in the home situation. "She tore her clothes in threads." Two lobotomies, August 24, 1944, and April 26, 1945, gave her a new start in life.



Figure 141 (c). Case 262. Eight months after standard lobotomy. She returned to her habit of smashing toys.

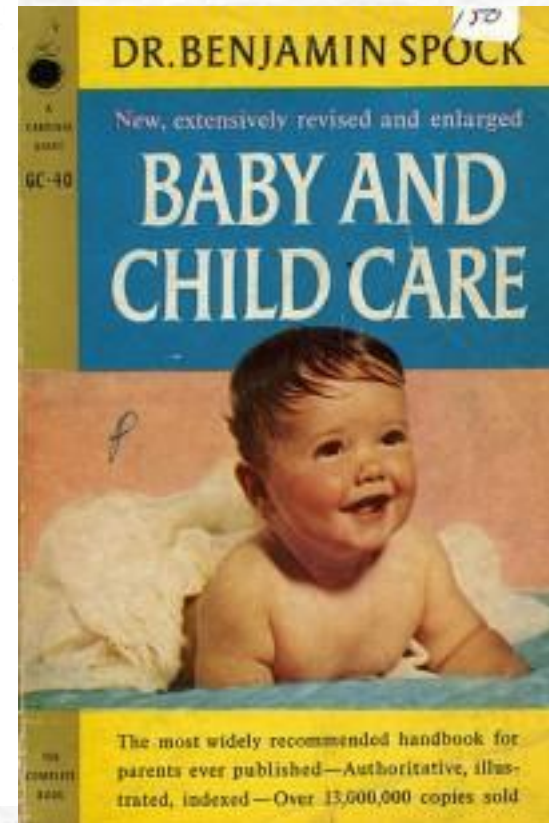
Hvor meget er evidensbaseret i kommunerne?



# Opinion based technologies

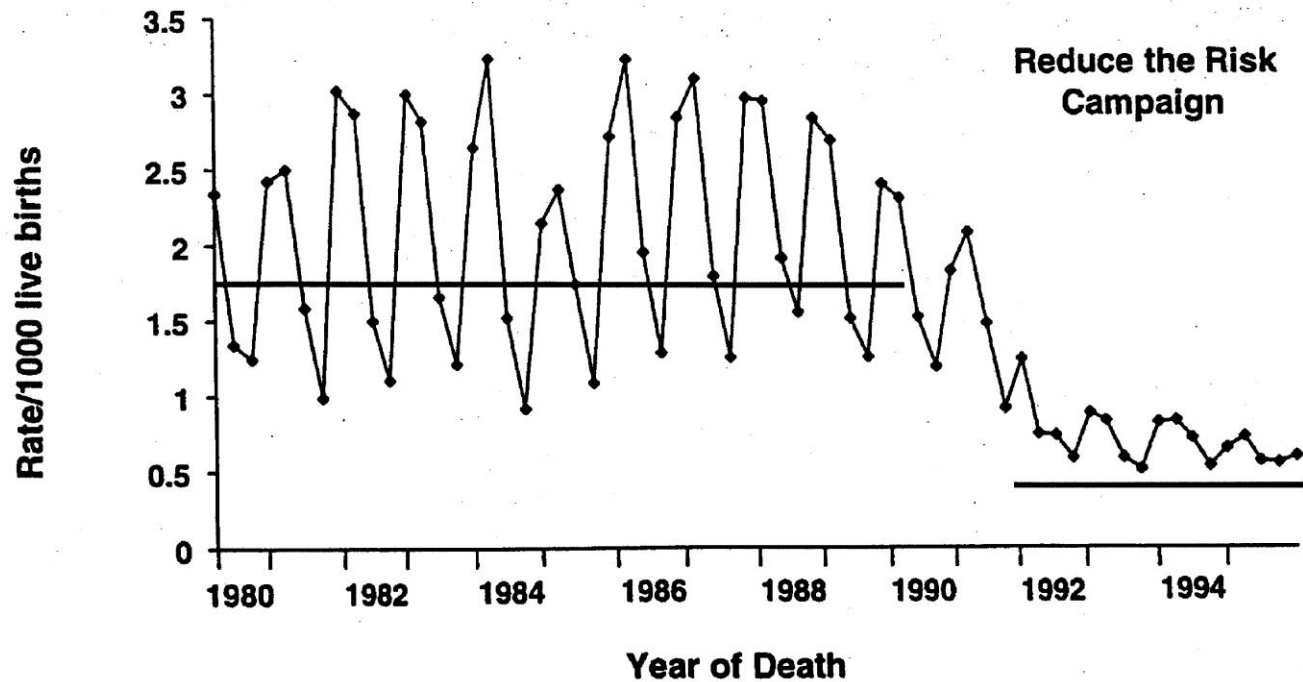
There are two disadvantages to a baby's sleeping on his back. If he vomits, he's more likely to choke on the vomitus. Also, he tends to keep his head turned toward the same side—usually toward the center of the room. This may flatten that side of his head. It won't hurt his brain, and the head will gradually straighten out, but it may take a couple of years. If you start early, you may be able to get him used to turning his head to both sides by putting his head where his feet were the time before, each time you put him to bed.

I think it is preferable to accustom a baby to sleeping on his stomach from the start if he is willing. He may change later when he learns to turn over.

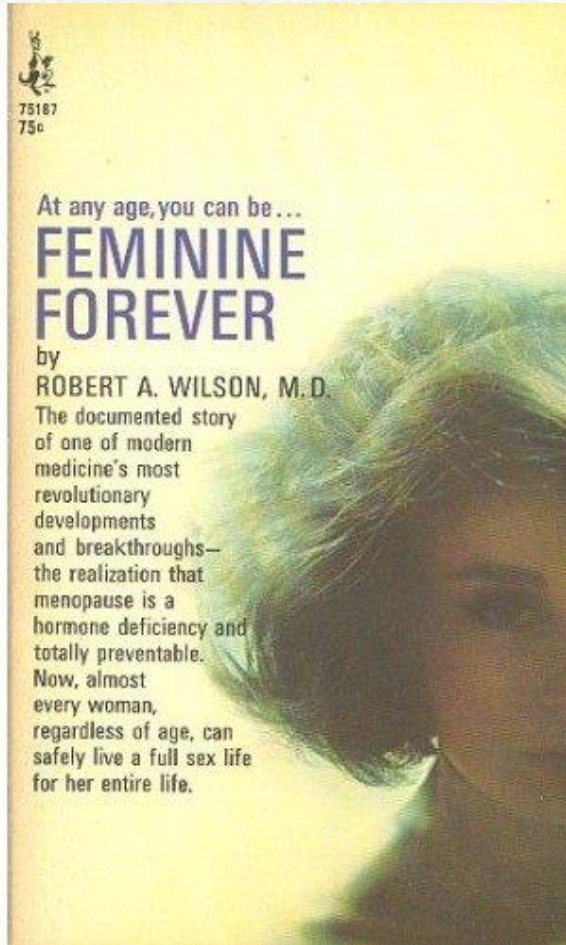




**Figure 3: Sudden Infant Death (SID) incidence (live birth to one year) by quarter. England and Wales 1980-1995**



# Og endnu et eksempel...



- 40 år med fejlbehandling af kvinder i overgangsalder pga økonomisk interesser fra medicinalindustrien
- Først i 2002 blev der publiceret RCT af hormonbehandling
- Resultat: ingen beskyttende effekt af behandling på hjertekarsygdomme, men øget forekomst af brystkræft

# General health checks in adults for reducing morbidity and mortality from disease



Sundhedsminister Nick Hækkerup (S) får opbakning fra 3F til sit forslag om at få socialt udsatte til et sundhedstjek. Flere forskere har dog kritiseret forslaget for ikke at være dokumenteret gavnligt. (August, 2014)

- ☞ Cochrane review
- ☞ 16 studies
- ☞ Conclusion:
  - Health checks did not reduce morbidity or mortality
  - Harmful outcomes poorly studied



# Improvement: 2009-2015

## QI:

- Lack of scientific knowledge about 'medical errors'
- Increasing interest due to financial crisis
- Consultancy firms with 'quick fix' solutions
- QI research: a strange mix of
  - Problems known from social science
  - Solutions inspired from health science





Spoon + fork = spork



# Patientsikkert Sygehus - kort fortalt

- Baggrund= 2 store udfordringer
  - mange patienter får ikke anbefalet behandling
  - implementering tager lang tid
    - => brug for nye metoder til understøttelse af best practice og til hurtig implementering
- Dansk Selskab for Patientsikkerhed, IHI, Trygf.
  - 5 deltagende sygehuse (Hillerød, Næstved, Horsens, Kolding, Thy-Mors)
  - 2010-13
  - 11 pakker
  - Budget 55 mio. kr.
  - 2018 deltagere i kompetenceudviklingsforløb
  - Evaluering af COWI 2014

# De 11 pakker

Intensiv Kirurgi Medicin Sengeafdeling **Alle**

AMI pakken 	CVK pakken 	Hjerte Insufficiens pakken 
Hjerte medicin 	KAD pakken 	Kirurgi pakken 
Medicin afstemning 	Mobilt akutsystem 	PVK pakken 
Respirator pakken 	Sopels pakken 	Trykstyr pakken 

# Metode

Forbedringsmodellen som lokalt implementeringsværktøj

Løbende lokal registrering af på forhånd fastlagte proces- og resultatindikatorer for hver pakke med brug af statistisk proces kontrol

Etablering af lokale teams til at drive implementeringen

Løbende udvikling af sygehusledelsens og personalets kompetencer ved faglig sparring og backup fra DSFP og internationale eksperter fra IHI ved

- › Læringsseminarer
- › Kurser, temadage og workshops
- › Sygehusbesøg med konkret sparring på de enkelte afdelinger
- › Telefonkonferencer



# Forbedringsmodellen - PDCA



# Hvad var formålet med PSS?

- ☞ At bidrage til løsningen af to udfordringer:
  - Ikke alle patienter får anbefalet behandling
  - Implementering af ny praksis tager (alt for) lang tid
- ☞ At afprøve en ny form for 'løsning'
  - At introducere 11 pakker (baseret på tjeklister )
  - At anvende model of improvement som metode
- ☞ Målet for Patientsikkert Sygehus blev formuleret kvantitativt som:
  - En overordnet effekt på HSMR & GTT
    - Mål for HSMR=reduktion på mindst 15%
    - Mål for GTT=reduktion på mindst 30%
  - Andre kvantitative mål for de enkelte pakker (som ikke indgår i eval., fordi der er 'evidens' for effekt i forvejen)

# Hvad sagde evalueringen?

- ☞ Overordnet effektmål kunne ikke vurderes
  - Ingen entydig udvikling i HSMR & GTT
  - Data ikke indsamlet til forskningsformål
- ☞ Konkrete proces- og resultatmål fra 33 afdelinger
  - Implementeringsgrad af pakker fra 12% til 62% (på udvalgte områder med bedst datagrundlag)
  - Effektdata fra SPC kan ikke 'be- eller afkræfte en statistisk signifikant samvariation mellem implementeringsgrad og behandlingsresultat'
    - Dog positiv tendens (12 afd positiv, 18 afd. neutral, 3 afd. negativ udvikling i resultatmål)



# Hvad sagde evaluering (2 af 2)

- ☞ Patientfokus og patientsikkerhedskultur
  - Fald i kulturens 'modenhed' jf Kristensen (men kan skyldes tidspunkt for måling og valg af instrument)
- ☞ Interviews (ca. 150 respondenter) tyder på 'anden effekt'
  - Karakteristika for det succesfulde sygehus = 'det hele hænger sammen', forankret på alle ledelsesniveauer og i kvalitetsorganisationen
  - Medarbejdertilfredshed, patientfokus og faglig stolthed
- ☞ Med andre ord:
  - Evalueringen er uklar mht effekt
  - Evalueringen søger desværre ikke svar på det oprindelige formål (om pakker + model of improvement er løsningen på udfordringerne i sygehusvæsenet), men har fokus på effekt

# Andre eksempler på effekt evaluering af komplekse interventioner uden klar konklusion



# Og hva' så?

- ☞ Den sædvanlige kritik af evalueringer
  - skulle være tænkt bedre ind fra starten
  - bedre data, længere tid, flere ressourcer
- ☞ Og alt det vi ikke fik evalueret
  - Fordele og ulemper ved pakker, tjeklister og konsulentdrevet undervisning
- ☞ Skulle man have lavet en anden form for evaluering?



# 1. What is a complex intervention?

*(Craig et al. Br Med J 2008;337:a1655.; the MRC guidance on complex interventions)*

## **Box 1 What makes an intervention complex?**

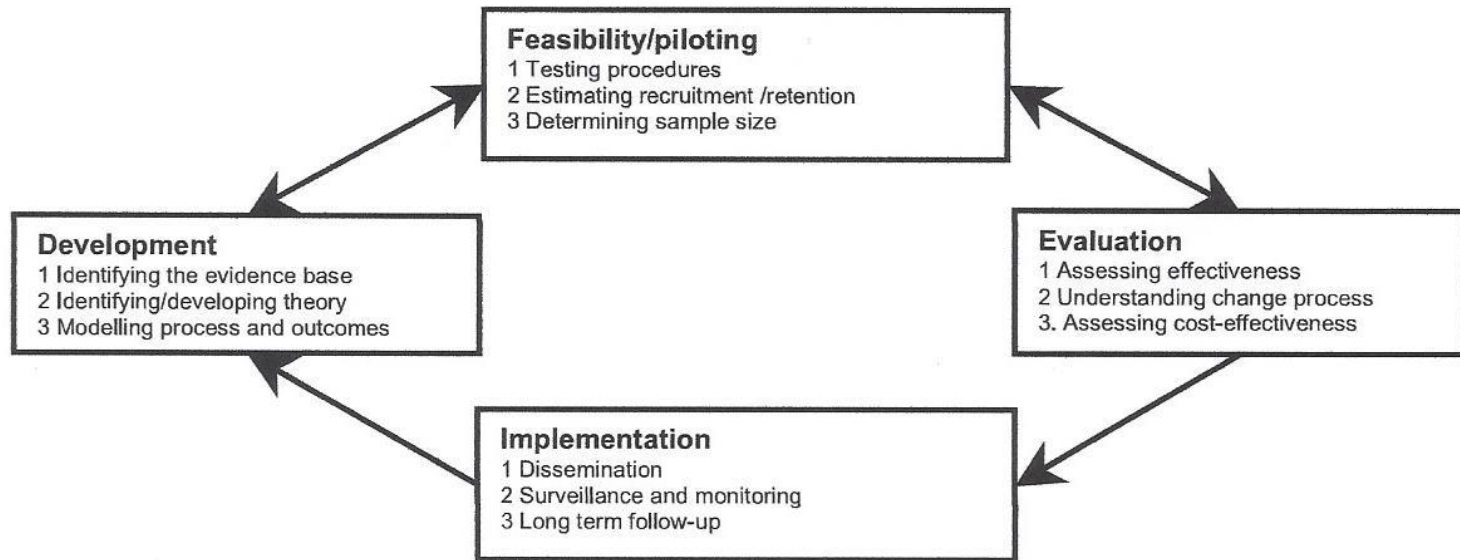
- Number of interacting components within the experimental and control interventions
- Number and difficulty of behaviours required by those delivering or receiving the intervention
- Number of groups or organisational levels targeted by the intervention
- Number and variability of outcomes
- Degree of flexibility or tailoring of the intervention permitted

## Also:

- Complex interventions are multi-level (policy, institutional, individual), multi-component, represents system thinking, always locally adapted, lack of linear, well-evidenced causal pathways
- Complex interventions are more than the sum of the parts; Hawe et al., BMJ 2004 (and that's why standardization have weak effect)

# MSC guidance

Figure 1 Key elements of the development and evaluation process



This is a SPORK!





# Eksempler på effekt-evaluering af komplekse interventioner med klar konklusion om effekt (og med brug af RCT og MSC model)

## **TeleCare Nord** Telemedicin i storskala



Overvåget læge    Sygeplejerske    Medicinske plejerske    Patient    Kommune

KRONIKER  
ENHEDEN  
NORDJYLLAND

### "Spør 2": Forløbskoordination

#### Puljen omkring forløbsprogrammer:

- Udvikling og afprøvning af modeller for forløbskoordination.

Udpejning af 3 kommunale deltagere og 3 sygehusedtagere

Gerne fra fora –

# Er Patientsikkert Sygehus en complex intervention?



Simpel: hvis det betragtes som 12 enkeltstående interventioner

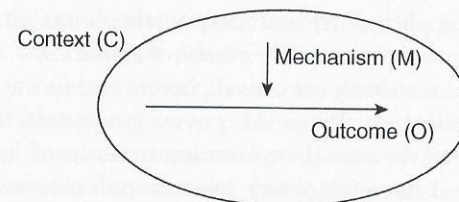
Kompleks: hvis det betragtes som en overordnet intervention til ændring af kultur på et sygehus

# The pragmatic definition of a complex intervention

An intervention is 'complex' when it is helpful to see, and analyze it as such (*Petticrew. EJPublic Health 2011;21;397*).



*Context (C) + Mechanism (M) = Outcome (O)*





# What do we want to know about QI?

1. Does it work? (Effectiveness)
2. Does it provide savings?
3. It is the best thing to do? (Cost-effectiveness?)
4. Does it work else-where?
5. How to spread?

# What do we want to know?



1. Effectiveness
2. Savings
3. Cost-effectiveness
4. Does it work else-where?
5. How to spread?

# Other goals in evaluations

- ☞ Was the project a success for us?
- ☞ Can we use the experience to improve beyond the project period?
- ☞ Can we document our progress
- ☞ Show political legitimacy
- ☞ Learning & employee satisfaction
- ☞ Promote public relations and awareness
- ☞ Add to the knowledge base (science)
- ☞ "prove we were right / we are the best"
- ☞ "If we evaluate then it is evidence-based"
- ☞ etc





**Evaluation**

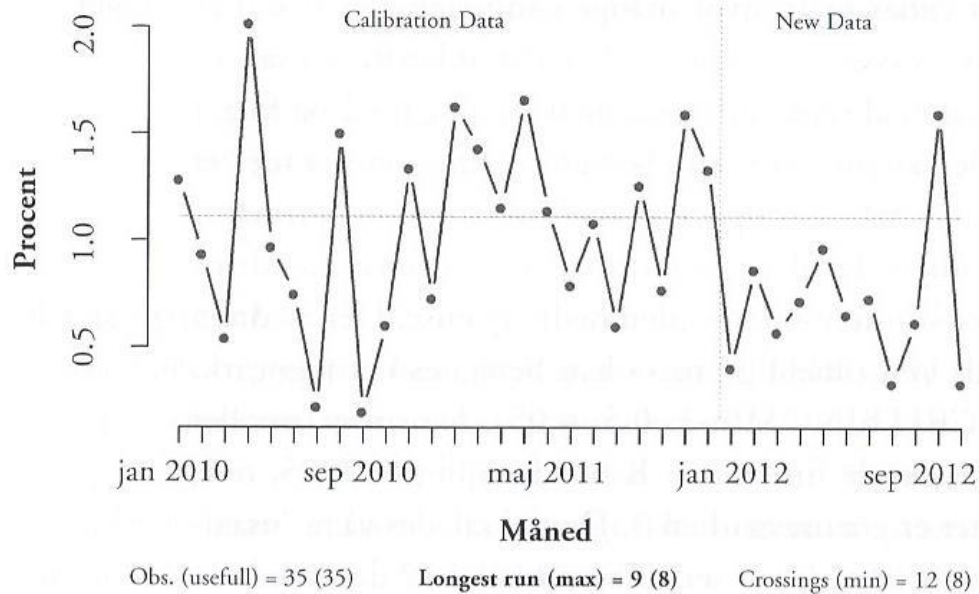
**Research**

Perhaps we should clearly distinguish between evaluation and research and make sure we do both each time?

# Example: Does it work?/ Is it effective?

(Anhøj, Kompendium i kvalitetsudvikling, 2015)

## Postoperativ mortalitet før og efter sikker kirurgi-tjekliste



**Figur 2.5:** Seriediagrammet viser et eksempel på anvendelse af principperne for brug af seriediagrammer til dokumentation af forbedringer: Indikatoren er den postoperative mortalitet på et dansk sygehus før og efter indførelse af WHO's sikker kirurgi-tjekliste. Medianen er beregnet for baselineperioden og forlænget ind i efter-perioden. Diagrammet viser et skift i efter-perioden, idet den længste serie består af ni datapunkter mod forventet højst otte. Skiftet går i den ønskede retning, og forandringen repræsenterer derfor en forbedring. Det næstsidste datapunkt ligger dog iøjnefaldende højt i forhold til de omgivende punkter, og man bør overveje, om dette repræsenterer en afvigelse, og i så tilfælde, om der findes særlige forhold, der kan forklare denne.

Vil svaret vedrørende 'effekt' være det samme i forskning og i evaluering?  
Svar: Nej

# 1. Does it work? (Research questions regarding effect is more demanding)

1. Does the effect depend on context? If yes, how?
2. Does it work compared to best practice or just compared to 'poor practice'?
3. Was it the best way to achieve the improvement? Or what other ways are there?
4. Does it work over a longer time period? And does it work elsewhere / or in other situations?
5. Does it work without side effects?
6. How big was the effect (quantitative)
  1. Before vs after the intervention
  2. Compared to the counterfactual
  3. Compared to best alternative interventions



# ... en bisætning om COWI's fortolkning af 'effekt'

- ☛ Pakkernes effekt er veldokumenteret internationalt
- ☛ Sygehusene har selv udarbejdet 'veldokumenterede forskningspublikationer' om pakkernes effekt i Patientsikkert Sygehus
- ☛ COWI; 2014

# ... en bisætning om Model of Improvement

- ☞ Plan-Do-Study-Act er klassisk målstyring
- ☞ Den statistisk test i modellen er det 'nye'
- ☞ Taylor et al. Systematic review of the application of the Plan-Do-Study-Act method to improve quality in healthcare. BMJ Qual Saf 2014.
  - Kun 2 ud af 73 studier bruger model of improvement stringent

Example: Are labour-intensive efforts to prevent pressure ulcers cost-effective? Mathiesen et al. J Med Eco 2013;16:1238-45.

- Valg af outcome mål problematisk
- Data utilstrækkelige (små serier, insignifikant, kortvarig)
- Ingen viden om counterfactual, alternativer eller 'best practice' på området
- Ingen viden om bivirkning ved pakker og tjeklister (eller hvad der nedprioriteres i stedet for)

Derfor kunne beregningerne ikke laves på empiriske erfaringer fra Patientsikkert Sygehus



# What factors should determine the choice of evaluation methods?

- Hvad er en evaluering?
- Definition: En omhyggelige tilbageskuende vurdering af værdien af ... (en intervention) ... med henblik på fremadrettet beslutningstagning og forbedring af praksis
- (frit oversat efter Vedung, 2009)

# The different types of evaluation

- **Goal achievement evaluation** ask questions about the extent to which goals have been achieved
- **Process evaluation** ask how the implementation of the intervention went
- **Interactive evaluation** ask questions about the stake holders' perception
- **User evaluation** focus on the users' experiences
- **Effect evaluation** focus on comparative outcomes
- **Economic evaluation** focus on comparative measures of both costs and outcomes
- **Programme (realistic) evaluation** focus on what works for us and why

According to evaluation theory the choice of evaluation method should depend on the purpose of the intervention

# What factors should determine the choice of research methods?

## • Type of research question? (Yes!)

Strategy	Form of Research Question	Requires Control of Behavioral Events?	Focuses on Contemporary Events?
Experiment	how, why?	Yes	Yes
Survey	who, what, where, how many, how much?	No	Yes
Archival analysis	who, what, where, how many, how much?	No	Yes/No
History	how, why?	No	No
Case study	how, why?	No	Yes

Figure 1.1 Relevant Situations for Different Research Strategies  
SOURCE: COSMOS Corporation.

## • Type of intervention ? (Not according to Petticrew. Men MSC model god til effekt-evaluering af komplekse interventioner)

## • Others?

- Yes, situation, time and money may restrict the set of possibilities i.e. only the second best may be realistic



# Final comments

- Perhaps in the future we should be better to:
  - Differentiate between evaluation and research – and do both
  - Don't do 'naive' evaluations of effect on organizational level without controls, but do RCT's when ever research questions on effect are vital
  - Use many other research designs than RCTs to answer other questions than 'comparative effect'
  - Use the university (social science) more (and build long term relationships with social science for education, evaluation and research purposes)
  - Don't believe in a universal design for evaluation or research for complex interventions, but
    - Research question should determine research design
    - Purpose of intervention should determine type of evaluation

# There's more than one SPORK!



Tak for opmærksomheden!

