The unbearable lightness of healthcare policy making: a description of a process aimed at giving it some weight

I B Scheel, K B Hagen, A D Oxman

Objectives: To investigate whether a structured process to involve policy makers in designing a research project on a return to work insurance policy would yield evidence that was relevant, useful, and used in policy decisions.

Study design: Case study.

Setting: Norway.

Participants: Two researchers from the National Institute of Public Health and four representatives from respectively the National Insurance Administration, Norwegian Confederation of Trade Unions, Confederation of Norwegian Business and Industry, and Norwegian Medical Association.

Intervention: Structured discussions of the research, including the objectives, interventions, design, and interpretation of the results.

Results: The participants succeeded in designing and completing a cluster randomised controlled trial through the participatory process. Intermediary results from the trial have been used in practical planning within the National Insurance Administration, but there are few indications that the main results of the trial have been used.

Conclusions: This approach of involving policy makers in the research planning process when political or organisational values are at stake did not succeed in this case. The salient explanations for this are conflicting interests of the organisations involved in the process and the research findings were in conflict with those interests.

Healthcare policies are interventions into peoples’ lives, and it is reasonable to require the same documentation of effects as one expects for healthcare interventions aimed at individuals. In fact, because the essence of policy decisions is to make choices for populations and thereby reduces the field of individual choice, it is possibly even more important that they are based on the best evidence available. However, policy is by nature rooted in political values and more often than not supported solely by pooled judgements and expert opinions, which make healthcare policy making seem unbearably light.

There is limited evidence of the effectiveness of strategies to improve the use of research evidence in policy decisions. However, it is commonly suggested that researchers and policy makers should communicate directly. Caplan and colleagues present the two community thesis or two culture thesis, where they describe the policy makers and the researchers as members of two cultures who lack the will or ability to relate to each others’ perspectives. In their systematic review of literature on policy makers’ perceptions of barriers to the use of research evidence, Innvaer and colleagues found that the most commonly mentioned barriers were mutual mistrust and the lack of personal contact, in addition to lack of timeliness, relevance, and availability. Granados and colleagues offer a suggestion on how to involve policy makers in the research process through a structured process.

We had the opportunity to use such a structured process when our research group in 1997 was asked to do some research on the implementation of Active Sick Leave (ASL); a new activity oriented sickness benefit arrangement offered by the National Insurance Administration (box 1). Though enthused by the prospects of addressing an important policy question in an evaluation, we were nevertheless concerned that the research we were to undertake would not be used by policy makers. Considering the increasing recognition of how stakeholders influence decision making processes, we engaged representatives from four stakeholder organisations (box 1) with powerful influence on policy decisions concerning work and health issues to take part in the research planning process. The principal aim was to ensure that the results of the forthcoming study, in fact, would provide answers to important policy questions. Furthermore, we anticipated that the participatory process would give the policy makers a sense of ownership and commitment to the evaluation and its results, and thereby promote use of the research results in practical policy decisions.

At the time we started our work with the group of policy makers, the policy makers were concerned with how to expand the use of ASL while we (the researchers) were also concerned with the lack of evidence to underpin such an expansion. The research results from the ASL trial are reported elsewhere. This article describes how researchers and policy makers worked together to design a robust evaluation study, and discusses whether this participatory process yielded research results relevant to and used in policy decisions.

METHODS

In addition to two researchers, the participants in our group were representatives of important national organisations; the National Insurance Administration (NIA), Norwegian Confederation of Trade Unions (LO), the Confederation of Norwegian Business and Industry (NHO), and the Norwegian Medical Association (Dnlf). All of these organisations are influential in the shaping of all national policies concerning workers’ absenteeism and sickness benefit insurance policies. In a series of meetings from January 1997 to September 1998, the four policy makers and two researchers attempted in a structured process to sequentially discuss and agree on a series of questions essential to the design of the research project. The questions were:
Active sick leave (ASL) is a Norwegian health insurance option that enables sick employees to return to modified duties at the workplace while the National Insurance Administration (NIA) continue to provide 100% of normal wages. For a sick worker to use the arrangement, some degree of communication and collaboration between the general practitioner, the employer, and the local insurance officer is required. The policy was introduced in 1993, but uptake of ASL was slow. In 1995 less than 1% of eligible workers were registered as ASL users despite a broad political will to expand the use of ASL. The politics underpinning ASL are called “The Workline” and is based on the idea that activity and work are better for people than idleness and isolation, given that some residual working capacity exists. The government’s policy is that all public social and health benefit schemes should be organised such that active, even if modified, participation in the job market will become the most attractive option for all.  

In 1994, four representatives from the National Insurance Administration, the Norwegian Confederation of Trade unions, the Confederation of Norwegian business and Industry, and the Norwegian Medical Association formed a group to endorse the use of ASL. In 1996 this group commissioned our initial study of barriers to the use of ASL in the communities, a project that was completed within the same year. The ASL trial, cluster randomised at community level followed as an extension of this work.  

1 What are the most important potential effects of active sick leave? (To decide on outcome measures)  
2 How can these potential effects be measured? (To decide on data collection)  
3 What are the best options for improving the use of active sick leave? (To decide on interventions)  
4 How can we evaluate whether those options are effective? (To decide on research design)  
5 How will the results of the evaluation be interpreted and used?  

We kept a record of the meetings and the outcomes of each discussion. The research design was a direct outcome of the structured process. To assess whether the research results were put to use in policy decisions we rely on our own observations and reflections of events linked to the ASL policy, the ASL trial, and the structured group process. Although we have discussed earlier versions of this manuscript with the people involved in this process, the events are described from the researchers’ perspective and therefore we cannot claim that the observations are entirely exhaustive or the reflections are entirely objective.  

RESULTS  
What are the most important effects of ASL?  

The aims of ASL were not explicit, and the group discussions revealed surprisingly inconsistent expectations of effects within the group. Some were “hard” effects that could be used as outcome measures, such as days off work and preventing drop out from the workforce, but many were process oriented and coloured by different and even opposing values. Examples of the latter include: to reinforce the sick employee’s sense of belonging, to utilise the remaining work capacity of people on sick leave, and to support worker-employer collaboration to improve the work environment.  

After several group meetings and still not reaching agreement, we explored the official documentation underpinning the policy. We found the following expected effects: reduced risk of long term disability, reduced average length of sick leave, and improved quality of life for the workers on sick leave. We agreed to base the outcome measures on those.  

Although the policy makers in our group believed that increased use of ASL would have a positive effect on all of these outcomes, it was agreed that ASL theoretically could have a negative impact on each of these.  

How can these results be measured?  

Once we agreed on the key effects, it was comparatively easy to agree on how these should be measured. The NIA routinely registers data for all workers on sick leave. These data could be used to measure long term disability and average length of sick leave. Long term disability was defined as sick leave lasting for more than one year. We chose patients on sick leave for back pain as the target group because they constitute a large proportion of total sickness absenteeism in Norway and because a probable effectiveness of early return to work interventions for this group was indicated in the literature.  

We selected the short form of the Rand health status measure (SF36) to measure health related quality of life.  

What are the best interventions to improve the use of ASL?  

For practical reasons the policy makers were most interested in an intervention that, if effective, could easily be introduced throughout the country and that could fit in with ordinary functions, thereby not requiring a special organised service. On the other hand, the literature on changing professional behaviour caused the researchers to opt for a more active and resource demanding intervention. The group therefore agreed to evaluate two implementation strategies as interventions to improve the use of ASL.  

How can we evaluate whether those options are effective?  

The policy makers meant it was important to evaluate the effects of the interventions in “the real world”, as compared with a laboratory setting, such as a single workplace. As expressed by one: “We don’t want an artificial setting where all the lessons learned are for the researchers and all the project’s structures collapse as soon as the researchers withdraw from the scene.” Another participant noted that the use of ASL requires the active participation of the general practitioner, the employer, and the local NIA officers, as well as the patient. This matched the researchers’ point of view. We agreed that the evaluation would have to be undertaken on a community rather than on an individual patient level.  

It was decided that a randomised controlled trial would offer the most rigorous evaluation as it would ensure a fair comparison by controlling for the many external factors that might effect the implementation of the interventions. To ensure that the effects would be evaluated in the real world, we randomised municipalities and targeted the interventions at all the key people involved in the use of ASL (general practitioners, NIA officers, employers, and people on sick leave with low back pain).  

How should the potential results of an evaluation be interpreted?  

This was by far the most difficult question to discuss. Multiple combinations of results were possible, each requiring explicit considerations of values and potential policy actions. The participants found it difficult to commit themselves. We discussed scenarios such as: What if the research results show that an increased use of ASL has significant effects on reducing the average number of days on sick leave, but also causes reduced quality of life for the people involved? Or, What if a larger proportion of people with back pain never return to work but the reported quality of life improved?  

Policy options included, for example, (a) to continue supporting the promotion of ASL to increase the use, (b) to leave ASL as an optional arrangement to people on sick leave, but not support promotion of the policy, or (c) to advocate the
withdrawal of the arrangement. We did not reach an agreement about the implications of possible results.

The direct output of the participatory structured process was the trial design and trial results, as schematically summarised in figure 1.

**EVENTS AFTER THE RELEASE OF THE INTERMEDIARY RESULTS**

The intermediary results of the ASL trial showed that the proactive intervention was effective in increasing the use of ASL, while there was no measurable difference between the passive intervention and the control communities. Releasing these results had immediate impact on the NIA. Managers were eager to learn from the trial and elements of the proactive intervention were incorporated in the organisational strategy plans.

However, at the same time, but at a different level, the stakeholder organisations represented in our discussion group came together to produce a plan to increase the use of ASL. The aim was to reduce workers sick leave by 20%. The elements of this plan were identical to the elements of the passive intervention in the ASL trial and ignored the fact that this intervention was not effective for increasing the use of ASL, that the effect of ASL on days off work was still not known, and the results of our trial for this outcome would soon be available.

At this point we were aware that we might have to deliver research results that were widely unpopular. Already before the results on the health outcomes were released, the NHO (employers’ association) medical advisor warned us that the organisation had decided to believe that ASL in fact is effective, regardless of what the study might show. The labour union was equally dismissive, based on the values underpinning the policy.

**EVENTS AFTER THE RELEASE OF THE RESULTS FOR HEALTH OUTCOMES**

In the end, the trial found that expanding the use of ASL did not have measurable effects on length of sick leave, disability,
DISCUSSION

We succeeded in planning and conducting a robust evaluation on the effects of promoting and expanding the use of ASL.15

The relevance, as well as the robustness of the results is, however, questioned by the same organisations that were represented in planning the trial and selecting the outcome measures, and the results of the trial have generally been ignored in successive policy actions.

One reason may be that ASL still is an attractive intervention that is consistent with the values of all the stakeholder organisations and authorities. As a concept, ASL is highly congruent with principal aims for the Work Environment Act, which is of prime importance to the labour union. The other side of the same coin is that it represents a low to no cost tool for employers to fulfill parts of their obligations to the Work Environment Act. The authorities aim to shift the use of resources from passive payment of social benefits to more active interventions for the beneficiaries. ASL is an intervention that supports this aim.

As an intervention representing the Workline (box 1), ASL has ideological power because of its potential to mask basic conflicts in perspectives.16 It has been suggested that policy makers should consider work as a fundamental right, also for people with diminished capacity, and those who think that stricter demands should be placed upon employees receiving benefits while still having some capacity for work. Acceptance of the ASL trial results might threaten this politically powerful alliance, and stakeholders with a vested interest in maintaining this union between traditionally opposing organisations might therefore ignore the research results.

However, in Norway sick leave and disability is an important economic and political burden and the authorities’ principal aims in promoting ASL are still to reduce these burdens. Many parties had put their faith in the unproved merits of ASL. The results of our trial, if accepted, suggest that a key tool to reduce sick leave and disability is unlikely to help achieve these aims. Lacking other plausible alternatives, policy makers might have felt compelled to act despite the new, discouraging evidence the ASL trial brought forward.

The results of our trial do not imply that ASL should not continue to be offered as an option to people on sick leave. Nor do they bring into question the underlying values of keeping people active and engaged in work. But they do suggest that ASL, at least as it is currently used, is disappointing with respect to its potential to reduce the huge burden of sick leave in Norway. Other, more effective strategies are needed to solve this problem.

There are possible limitations to the structured process described in this paper. There is little doubt that the most important stakeholder organisations were represented in our group but, as one participant suggested, each representative’s level of influence in the respective organisation might have been too low to set off diffusion of the idea of evidence based policy making. Continuity was also somewhat disturbed by replacement of several participants during the lengthy process.

Another constraint may be that not only the policy makers, but also the researchers held stakes in this process. Initially, the policy makers approached us with an explicit research question, appropriate funding, and a non-committal attitude to the forthcoming research. The need to secure continued funding for the ASL trial might have caused some cautiousness in the process on our part, especially in the timing of discussing potentially undesirable results.

Although the paper has been thoroughly discussed with the participants in the structured process and has been rewritten several times, the case study is based upon the researchers perception of the events, and only those events we had access to.

Innvaer and colleagues discuss the use of research results among policy makers in terms of three categories.17 Direct use indicates the application of research results that are relevant for a solution. Enlightening use helps to enrich and deepen understanding of the complexity of problems and unintended consequences of action. Selective use is strategic and involves use of research to legitimate and sustain predetermined positions. Despite the participatory process to promote the use of relevant research results in practical policy making, this study describes a case of selective use of research results, where
policy makers choose to ignore results that do not support their predetermined positions.

ACKNOWLEDGEMENTS
We thank the organisation representatives from the National Insurance Administration (NIA), Norwegian Confederation of Trade Unions (LO), the Confederation of Norwegian Business and Industry (NHO), and the Norwegian Medical Association (Dnlf), who participated in the structured discussion sessions and were instrumental in ensuring a rigorous evaluation. We also thank their respective organisations for co-funding the Active Sick Leave Project along with the Norwegian Ministry of Health and Social Affairs.

Authors’ affiliations
I B Scheel, K Birger Hagen, Department of Social Services Research, Norwegian Directorate for Health and Social Affairs, Oslo, Norway
A D Oxman, Department of Health Services Research, Norwegian Directorate for Health and Social Affairs

The views expressed in this paper do not necessarily represent the views of the Norwegian Directorate for Health and Social Services.

REFERENCES
15 Sundberg A. Aktiv sykemelding for et inkluderende arbeidsliv. Dagens Næringsliv 2002;Feb 14.