Economic Analysis

ECONOMIC OBSERVATIONS CONCERNING OPTIMAL PREVENTION AND COMPENSATION OF DAMAGE CAUSED BY MEDICAL MALPRACTICE

Michael Faure*

I. Introduction

In the economic analysis of law, damage is defined as an externality. The basic idea, therefore, is that legal rules should give the health care providers, the physician or the hospital, incentives to invest in care in order to prevent damage. One of the important questions in law and economics is, therefore, through what kind of legal rules the health care providers can be given appropriate incentives for optimal care, and thus to internalise the risk for medical malpractice. This prevention of medical malpractice can be achieved either through liability rules or through regulation. Possibly, other alternatives might fit better in order to avoid or lower the risk of medical malpractice as well.

The answer to these questions of course depends on the applicable legal framework. A traditional legal question is, more specifically, whether there exists an agreement based on a contract between the patient (the victim) and the health care provider (the injurer) causing the damage. If there exists a certain contract, as is presumed in most legal systems, this contract will also determine under which conditions the injurer is liable, and thus the contract forms the foundation for (any) compensation of the victim. However, this paper will not primarily deal with the question if medical liability is based on contract law or liability law, because the answer to this question depends on the actual facts of the case. Besides, it depends on the question if the physician himself (or a member of the hospital staff) or the hospital is liable towards a patient by contractual agreement. Moreover, the questions whether the liability of the health care provider is based on contract or tort does not seem to play a major role in legal practice any longer. However, as we will argue below, the existence of a contractual relationship may be important from an economic per-

* In cooperation with Aimée Verhulstomck.
spective since, in that case, the health care provider may be able to pass on the increased costs of liability via the price-mechanism.

2 The purpose of this introductory paper is to examine which legal instruments can be used to prevent medical malpractice (prevention) and how possible damage due to this medical malpractice can be compensated in an optimal way (compensation), both from the perspective of law and economics. This paper largely draws on earlier law and economics research for the European Group on Tort Law, for instance with respect to environmental liability\(^1\) but also on medical malpractice\(^2\). For the reader who is not aware of this earlier literature, the general principles of the economic analysis of tort, insurance and compensation are explained rather in detail; afterwards applications to the issue of medical malpractice are presented.

3 First, in a general overview the economic model of tort law is explained. In this section the paper tries to answer the question for what purpose liability rules are used and how liability rules work, especially focused on the medical field. Next the Coase Theorem is explained and applied to medical practice. Second, we will take a look at why optimal care is not only established through the liability rules, but why it is also necessary to have additional legal instruments, for example regulation, to reach the goal of optimal prevention. Thirdly, we will examine the issue of optimal compensation of damage caused through medical malpractice from an economic perspective. In addition to addressing the functioning of liability rules, we will also analyse which other models exist, showing alternatives for the liability model mentioned above. We will examine alternatives such as first and third party insurance and compensation funds and analyse how these focus on an optimal compensation to be awarded when damage occurs. The paper will conclude with a few directions for further research. These provide some conclusions of the economic analysis provided in this chapter which may be interesting for the other contributions.

II. Prevention

A. Goals of Tort Law from a Law and Economics Perspective\(^3\)

1. Prevention and Compensation: General Notions

4 The economic analysis of law in general and of tort law more specifically starts from the belief that a legal rule and more particular a finding of liability


\(^3\) This part contains first a general overview of the economic analysis of tort law and accident law in general and afterwards an application on medical malpractice. See for such an overview also M. Faure, Strict liability: economic analysis, in: B.A. Koch/H. Koziol (eds.), \textit{Unification of Tort Law: Strict Liability} (2002), 361–394.
will give incentives to potential parties in a tort setting for careful behaviour⁴. Thus, economists tend to stress the deterrent function of tort law. Lawyers on the other hand mention this deterrent function sometimes as well, but tend to attach more value to the compensation goal of tort law. This “victim protection”-argument is discussed in the law and economics literature as well⁵. In that respect it is, however, often stressed that the best way of victim protection is to avoid victimisation in the first place. Of course, no one will argue that prevention of accidents is not a way of victim protection as well. This difference in accents between both approaches is also characterised as an ex ante versus an ex post vision. Whereas lawyers tend to be more interested in the tort problem ex post, where there is a victim who needs to be compensated⁶, economists look at the tort problem in an ex ante perspective by asking the question how an ex post finding of liability will influence ex ante the incentives of potential parties in a tort setting for care-taking.

Of course the differences in approaches between lawyers and economists are not really that black and white. There are lawyers that stress the deterrent function of tort law as well⁷. And some economists pay attention to compensation issues by stressing that tort law should also aim at an equitable loss spreading⁸. One advantage of the economic approach is that the deterrent function and compensation goal are carefully distinguished so that the influence of various legal mechanisms that one would choose can be evaluated both with respect to the prevention and with respect to the compensation issue. The first scholar to analyse these problems from a law and economics perspective was probably the lawyer (!) Guido Calabresi from Yale Law School⁹. In his well-known book *The Costs of Accidents* Calabresi makes a careful distinction between primary, secondary and tertiary accident costs¹⁰. Primary accident costs are the costs of accident avoidance and the damage that finally oc-

---


⁵ Schwartz showed that rules of tort law may serve both the aims of deterrence and corrective justice (G. Schwartz, Mixed Theories of Tort Law: Affirming both Deterrence and Corrective Justice, [1997] *Texas Law Review* (TLR) 75, 1801–1834).

⁶ This tendency is partially explainable through the case method used in common law countries for law teaching; it focuses very much on solving one particular case, sometimes neglecting the consequences of the solution in one case for the behaviour of other parties and thus of safety in general.

⁷ For instance in Austrian tort law prevention is clearly stressed as one of the functions of tort law as well.


curs; secondary costs refer to the equitable loss spreading and tertiary costs are the costs of administering the legal system. Tort law should give incentives to reduce the total social costs of accidents. Thus the central goal of tort law was given: it should give incentive for a minimisation of accident costs. This notion of Calabresi was taken up later by economists who have formalised this issue\textsuperscript{11}. The first goal of tort law is the minimisation of primary accident costs: the costs of accident avoidance and the expected damage. Indeed, from a social point of view accidents do not only cause costs from the moment an accident occurs and harm is suffered; potential parties in an accident setting, both injurers and victims, make investments in care to avoid the occurrence of an accident.

2. Optimal Liability Rules: Basic Ideas and Models

a) Optimal Care Levels

Economists use classic cost/benefit analysis to determine what the level of care is that will lead to such a minimisation of the social costs of accidents. Not surprisingly this can be found where the marginal costs of care-taking equal the marginal benefits in accident reduction\textsuperscript{15}. Indeed, since care-taking has its price as well a legal rule should not give incentives to avoid every possible accident that could occur, but only damage that could be avoided by investments in care of which the marginal costs are lower than or equal to the marginal benefits in accident reduction. It might well be that extremely high care could additionally contribute to a reduction of the accident risk but the marginal costs of care-taking in that case might well be much higher than the additional benefit in accident reduction. Investments in care would in that case be inefficient and scarce resources would be spoiled\textsuperscript{15}. These levels of care where marginal costs of care-taking equal marginal benefits in accident reduction are referred to in the literature as optimal or efficient care levels\textsuperscript{16}.

b) The Economic Model

Tort law is thus seen as an instrument to deter activities worth being avoided on efficiency grounds through liability rules\textsuperscript{15}. The expectation to be held lia-


\textsuperscript{15} This distinction has been made by S. Shavell, \textit{Economic Analysis of Accident Law} (1987), 7.

\textsuperscript{16} This finding only holds in a risk neutral setting. In case of risk aversion higher investments in care might well be efficient since a risk reduction of accident will in that case also remove the disutility of risk from a risk averse person.


\textsuperscript{16} Although in this paper the preventive function of tort law is stressed, we already noticed that Calabresi has pointed at the fact that liability rules may equally aim at loss spreading (Calabresi’s secondary costs) and at a reduction of administrative costs (Calabresi’s tertiary costs). See G. Calabresi (supra fn. 10).
ble ex post should induce parties ex ante to take care or change the activity level in view of reducing the accident risk\textsuperscript{16}. Liability rules should, according to this economic model, thus be used to establish the efficient solution. Economists state that the goal of tort law is to minimise the sum of accident costs and the costs of accident avoidance\textsuperscript{17}. The sum, called the social costs of accidents, can be presented as follows\textsuperscript{18}:

\[ C = p(xy)L + A(x) + B(y), \text{ where:} \]

\[ C = \text{the sum of expected accident costs and the costs of care,} \]

\[ A = \text{the victim,} \]

\[ B = \text{the injurer,} \]

\[ x = \text{level of care of the victim,} \]

\[ y = \text{level of care of the injurer,} \]

\[ p = \text{probability that an accident will occur,} \]

\[ L = \text{magnitude of the loss.} \]

It is assumed that both parties are risk neutral, that the magnitude of the loss \((L)\) is independent of the level of care, that more care will reduce the probability \((p)\) of an accident and that only the victim \((A)\) suffers a loss. To minimise the social costs \((C)\), the levels of care must be set at \(x = x^*\) for the victim and \(y = y^*\) for the injurer. At these efficient levels, the marginal benefits from an increase in care (reduction of \(p(xy)L\)) equal the marginal cost of greater care\textsuperscript{19}. From an economic point of view, optimal care is thus not equal to the highest care possible. The highest care might lead to spillage, since the marginal costs would be greater than the marginal benefit in reducing the expected loss. Calabresi formulated this as follows:

"Our society is not committed to preserving life at any cost. In its broadest sense, the rather unpleasant notion that we are willing to destroy lives should be obvious. Wars are fought... Ventures are undertaken that, statistically at least, are certain to cost lives. We take planes and cars rather than safer, slower means of travel. And perhaps most telling, we use relatively safe equipment rather than the safest imaginable, because — and it is not a bad reason — the safest costs too much"\textsuperscript{20}.

The difficulty is therefore to find the efficient levels of care \(x^*\) and \(y^*\). In a tort situation, transaction costs are mostly prohibitive, meaning that the efficient

\textsuperscript{17} See G. Calabresi (supra fn. 10). See also S. Shavell (supra fn. 12), 5–6.
\textsuperscript{18} This is the basic model presented by S. Shavell, [1980] JLS, 1–25.
\textsuperscript{19} J.P. Brown, [1973] JLS 2, 323–349.
\textsuperscript{20} G. Calabresi (supra fn. 10), 17.
levels \( x^* \) and \( y^* \) cannot be reached through voluntary negotiations\(^2\). Therefore, the legal system should intervene to provide liability rules that will give incentives to the injurer that would lead to the efficient level of care. Let us now examine which legal rules, notably liability rules, would give these incentives for optimal prevention.

3. Negligence and Strict Liability

a) Negligence Rule and Optimal Incentives

If there were no liability at all, clearly the injurer would have no incentive for care-taking; therefore in a no-liability situation the externality (= possible damage) will not be internalised and an inefficient outcome will follow. One rule which may give incentives to the injurer to follow the efficient level of care is the negligence rule. The negligence rule is defined as a rule according to which the injurer will only have to bear the loss if he uses less than a legally required level of care, referred to as the due care level. If a negligence rule is adopted, the injurer will take optimal care, provided the due care required in the legal system is equal to the optimal care as defined in the model\(^2\). The negligence rule, as defined here, means that the injurer will be held liable if he spends less than the due care level required by the legal system, in other words if he acted wrongfully. Thus he will have to spend care to avoid the accident, but if he does so, he can avoid paying the expected damages. Of course the injurer could take more care than the legal system requires him to do under a negligence rule, but he will have no incentive to do so since he can already escape liability by following the due care standard. Of course the injurer could also spend less on care than the legal system requires him to. In that case he will have lower costs of care-taking, but he will have to pay damages in case an accident occurs. Since the optimal care standard was defined as exactly that level of care where the marginal costs of care equal the marginal benefits in accident reduction, taking less than the due care standard will not be interesting for the individual injurer since it will increase his total expected costs. Thus a negligence rule will lead to an efficient outcome as long as the legal system defines the due care as equal to the optimal care of the model.

b) Economic Model of Negligence Rule

This outcome can also be illustrated as follows. Assuming that the due care level required by the legal system is equal to the optimal level of care \( (y^*) \), the injurer will always follow the optimal care level. This is indeed the cheapest solution for him. If the injurer spent less than \( y^* \) on care, his total costs would be equal to

\(^2\) Otherwise the Coase theorem, which we will discuss later in this paper, would teach us that the efficient outcome would automatically follow through voluntary negotiations (R.H. Coase, The Problem of Social Cost, [1960] Journal of Law and Economics (JLE) 1, 1-44). However, as we will argue below, the Coase theorem may play some role in the context of medical malpractice.

\[ p(\alpha y^*L) + B(y^*) \]

If, on the other hand, he takes the efficient level of care, he will not have to bear the expected loss. Hence, the expected costs of the injurer are in that case only his costs of taking efficient care: \( B(y^*) \). The question whether \( B \) will take efficient care or not will, therefore, depend upon:

\[ p(\alpha y^*L) + B(y^*) \geq B(y^*) \]

If the sum of expected loss and the actual costs of care are higher than the costs of efficient care \( (y^*) \), the injurer will take efficient care. If, on the other hand, these costs were lower than the costs of taking efficient care, it would be cheaper for the injurer not to take efficient care. However, since we defined \( y^* \) as the point where the social costs are as low as possible, \( p(\alpha y^*L) + B(y^*) \) will always be higher than the costs of taking \( y^* \). Hence, the simple conclusion is that under a negligence rule, the injurer will always have an incentive to spend \( y^* \) on care and an efficient outcome will be reached. This is true as long as the legal system defines \( y^* \) as due care.\(^{23}\)

It should be stressed that the analysis presented here assumed that the injurer has money at stake to pay compensation to the victim and that only one party (referred to as the “injurer”) influences the accident risk. The influence of the victim on the accident risk is therefore assumed to be zero in the analysis presented here.

c) Strict Liability\(^{24}\)

A strict liability rule will also lead to the optimum. The reasoning is quite easy. A strict liability rule basically means that the injurer has to compensate in any case no matter what care he took. It is sometimes argued that this will lead the injurer to take excessive precautions or to take no care at all since he is liable anyway. Neither of these statements seems true. By making the injurer strictly liable, the social decision is in fact shifted to the injurer. It means that he has to bear all the social costs of accidents, being his own cost of care taking and the expected damage.\(^{25}\) Therefore, he will take exactly the same decision, being to minimise his total expected accident costs. We discussed in the model that this could be reached at the optimal care level. Therefore, the injurer will take optimal care since this is the way to minimise his total expected costs. Spending more on care would increase his costs of care-taking inefficiently and spending less on care would increase the expected damage inefficiently. Hence, strict liability provides optimal incentives to take efficient care as well.

\(^{23}\) See S. Shavell (supra fn. 12), 14.

\(^{24}\) This part is largely based on M. Faure (supra fn. 3).

4. Grey Areas Between Strict Liability and Negligence

a) Introduction

16 So far we assumed that there is a clear dividing line between fault liability (negligence) and strict liability but one can wonder whether in legal practice there are grey areas in between. Traditional economic analysis of tort law makes a clear distinction between on the one hand a fault liability, which in the American economic legal doctrine is usually referred to as negligence on the one hand and strict liability on the other hand. In that literature, of which we have summarised the main result, strict liability is usually relatively simply defined as a legal rule that forces the injurer to compensate the total losses of the victim irrespective of the behaviour of the injurer. Economic analysis has also paid attention to the fact that there are a number of situations which are formally not considered strict liability, but which amount to a strict liability rule, this is to say that it has the same effects.

b) Injurers with Lower Qualifications

17 A first strict liability situation may arise in situations where there is formally a negligence rule and has to do with the fact that the law usually defines the care required from an injurer in abstracto. From an economic point of view the optimal care, which the injurer should take to prevent an accident, will differ, taking into account the individual abilities of an injurer to avoid the accident. In other words: since the costs of taking care will differ, so will the optimal care. However, it may not always be possible for the legal system to calculate the actual costs of care of every injurer. The reason is that the costs of administering such a system of detailed classification will often be too high. However, an inevitable consequence of the fact that the legal system will require an average level of care and judge the care in abstracto (often referring to a bonus pater familias-standard) will be that some injurers may not have the capacity to follow this average done care standard, required by the legal system, since they could only reach this standard by investing very highly in prevention costs.

18 The result is that for some extremely low qualified injurers it may be too costly to exercise the average care required by the legal system under negligence; they will run the risk of being held liable and for them the negligence rule in fact amounts to a strict liability situation. This is a situation where we formally have a negligence rule, but where the injurer with low qualifications can never meet the average standard required by the legal system. For his case negligence in fact becomes strict liability. This situation may well be efficient,

27 The term bonus pater familias is a Roman based concept. It is often used in legal systems where it is called “the reasonable man standard”, and it refers to the conduct of a normally careful and prudent person, capable of conducting himself with care and diligence.
28 See S. Shavell (supra fn. 12), 75.
since it may give this lower qualified injurer an incentive to an activity level change\textsuperscript{29}. In other words: holding him to the high standard of care will induce him to refrain from the activity that for him apparently constitutes a too high risk. The question whether the standard of care should be applied in the abstract or in concrete way has its importance as well for the professional liability, as we will show below.

c) Professional Liability

The standard of care is generally established “in the abstract”. An appreciation “in the concrete” which would take circumstances such as age, education or intelligence of the injurer into account is very costly. However, if a classification can be done at relatively low costs, one can also notice this in legal systems, such as for example in Belgian tort law. Professional liability offers several examples. If, for instance, a doctor’s conduct will have to be examined, the criterion will not be the conduct of a careful and prudent person, but of a normally careful doctor. The average y\textsuperscript{*} standard (which is the efficient care level set by the legal system at an average level, because detailed classification of the due care standard of an injurer separately is too costly) will not be the average standard of a citizen, but the average standard of doctor. This will give the professional the appropriate incentives to take efficient care. In this case classification is obviously also not very costly, since it is clear that a professional can take care at lower costs than a normal citizen as far as it concerns his professional activity. Thus Dalcq writes concerning the liability of lawyers:

“La responsabilité professionnelle des avocats obéit aux règles de droit commun en ce sens que toute faute que n’aurait pas commise un avocat normalement prudent et avisé est susceptible d’engager la responsabilité”\textsuperscript{30}.

d) Breach of a Statutory Duty

Another example of a “grey area” is the situation where negligence is established on the basis of a breach of a statutory duty. Indeed, as we will elaborate below, very often the preventive measures to be taken by an injurer are also established by the regulator. This then leads to a combination of liability rules and regulation\textsuperscript{31}. The question then arises whether the violation of a regulatory norm should automatically be considered a fault, such as much legal systems do.

Since the regulator cannot identify atypical parties, a single regulatory standard will be used. Hence, a breach of a regulatory duty will automatically be considered a fault\textsuperscript{32}. Again, for individuals with lower capacities, for whom

\textsuperscript{31} See further on this combination below, no. 85 et seq.
\textsuperscript{32} The rationale for this rule will be further discussed below in no. 85 et seq.
meeting the regulatory standard will be impossible, this requirement (that the regulatory standard must be followed) may make negligence into a strict liability rule. Indeed, as soon as a statutory duty is breached, negligence will be considered as given without a further consideration of the possibilities of precaution of this particular injurer. This application of the negligence rule therefore, again, shows that there are grey areas between negligence and strict liability.

e) Reversal of the Burden of Proof

22 The last example of this “grey area” is again an issue which may play an important role in medical malpractice, being the reversal of the burden of proof of appropriate care. The issue of the economic meaning of the reversal of the burden of proof has been discussed in the context of a draft of a directive on the liability of suppliers of services which was presented a few years ago33 and was subsequently withdrawn34. Curran provided an economic analysis of the liability for supplier of services35 and paid much attention to the reversal of the burden of proof. Curran concludes that the reversal of the burden of proof will have no influence on the care to be taken by the supplier of services36.

23 This corresponds with the conclusion of the general economic model of tort law, being that both negligence and strict liability may give appropriate incentives for care to the potential injurer. However, a reversal of the burden of proof amounts to a shift in the risk of proof. Proving negligence amounts to costs for the victim, which are now shifted to the injurer, who then has to prove that he took appropriate precautions. If the injurer fails to provide this proof, he will have to compensate his victim. In that particular case the reversal of the burden of proof (under negligence) has in fact amounted to a strict liability situation37. The reversal of the burden of proof therefore constitutes indeed another example of a grey area between negligence and strict liability.

24 We should note, however, that in this context we only discuss the reversal of the burden of proving negligence (or appropriate care). The reversal may lead to a strict liability situation, but can lead to appropriate incentives and may therefore be efficient. This should be distinguished from another tendency in

33 Commission proposal of 9 November 1990, OJ, C 128 of 18 January of 1991; an earlier draft had been launched on 8 November 1989.
36 One should note, however, that this concerns a particular situation, given the contractual relationship between the provider of the service and the client. Hence, any agreement on the quality of the services may be passed on via the price mechanism.
the law to shift the burden of proof in case of causal uncertainty to manufacturers or providers of services. Reversal of the burden of proving causation may then lead to a situation that the manufacturer or provider may have to prove that his activity did not cause a certain damage. If he cannot meet this burden, this may amount to a situation whereby the injurer (and his insurer) are held to compensate losses which they have never caused. This reaction to causal uncertainty can hardly be considered as efficient34.

Let us now, before applying this general economic model of liability to medical malpractice, address once more the principle difference between negligence and strict liability.

B. Negligence versus Strict Liability: A Few Refinements

1. Information Differences

There is an important aspect of the difference between negligence and strict liability that should be mentioned, which may also have its importance for medical malpractice. This concerns the fact that the application of negligence requires high information costs from the judge who will have to set the due care standard. The information necessary to weigh costs and benefits and fix the optimal care may not be readily available for the judge. Strict liability shifts all costs to the injurer, who will then have to define the optimal care level. If one therefore assumes that the information on optimal precaution is better available to physicians than to judges, this constitutes an argument for strict liability. This information advantage may therefore constitute an argument in favour of strict liability for medical malpractice. One should, however, remember that this finding only holds in the models, such as the one that is e.g. developed by Shavell, which start from an assumption of risk neutrality. If risk aversion is introduced and the potential injurer is risk averse, strict liability is only efficient if in some way risk can be removed from the risk averse injurer, e.g. through insurance37. Moreover, we assume that the judge has accurate information on the amount of the damage. If courts err in assessing damages, strict liability will lead to underdeterrence40.

2. Insolvency of the Injurer

The seeming advantage of strict liability should be somewhat balanced, though. First of all, it has been assumed until now that the injurer has money

34 From an economic point of view a proportionate liability rule is a more adequate reaction. See for a further discussion of the potentially adverse consequences of a reversal of the burden of proof in case of causal uncertainty, M. Faure/T. Hartlieb, Remedies for expanding liability, [1998] Oxford Journal of Legal Studies (OJLS) 18, 681–706.


40 If, in other words, courts can more easily observe the socially desirable level of precaution than the exact amount of external harm, a negligence rule should be favoured. This point has been made by R. Cooter, Prices and Sanctions, [1984] Columbia Law Review (CLR) 84, 1343–1523.
at stake to pay compensation to the victim. If, however, the amount of the damage exceeds the injurer's wealth, a problem of underdeterrence will arise. Under strict liability, the injurer will consider the accident as one which is equal to his total wealth and will therefore only take the care necessary to avoid an accident with a magnitude equal to his total wealth. If that wealth is lower than the magnitude of an accident, he will take less than the optimal care and therefore a problem of underdeterrence arises under strict liability. Insolvency is less of a problem under negligence since under that rule the injurer will still have an incentive to take the care required by the legal system as long as the costs of taking care are less than his individual wealth. Taking due care remains indeed a way for the injurer to avoid having to pay compensation to the victim. Therefore the negligence rule can equally lead to an efficient result. This leads to a conclusion that will surprise many lawyers. Although high-expected accident costs are generally considered an argument in favour of strict liability, a fault rule might provide better results if the loss is of such a magnitude that it exceeds the injurer's assets. Therefore, the argument of increased losses, considered separately, does not seem sufficient for the introduction of strict liability rules. This is a simple application of the principle that the deterrent effect of tort liability only works if the injurer has assets to pay for the damage he causes. If an injurer is protected against such liability, a problem of underdeterrence arises.

28 This result is important since it shows that it would be erroneous to blindly promote the introduction of a strict liability rule under the argument that it would give better incentives for prevention to the injurer. Most legal systems still adopt a negligence regime for medical malpractice and this makes sense from an economic point of view. If this rule is applied correctly by the courts, a negligence system can give appropriate incentives for prevention to health care providers.

3. Positive Externalities

29 There is, in addition, another issue that has been mentioned in the literature that may balance the arguments in favour of strict liability. Gilead rightly pointed out that many activities do not only externalise harmful effects, but also have positive effects, which may be a reason to be somewhat cautious with strict liability. That is a consequence of the fact that strict liability incorporates activity level changes, whereas a negligence rule does not. It is not always obvious that increases in activity levels are socially undesirable.

---

61 See S. Shavell (supra fn. 12), 8.
Think about the manufacture of drugs or, in this context, the provision of health care. Most of these activities create huge social benefits. One should therefore be careful with the argument that strict liability is needed to internalise risk and control activity levels if the benefits of an increased activity were neglected.

There is finally, one other nuance, which has to be added. So far we assumed a strict liability rule whereby the injurer is “merely” held to compensate the damage he caused to his victim, not less but also not more. Problems may arise if strict liability is combined with other features, which may expand the burden of liability on physicians/hospital (staff). One could think of a removal of the burden of proving causation\textsuperscript{47}, joint and several liability and high (punitive) damages for (non-pecuniary) losses. These are formally different issues than the simple choice between negligence and strict liability. However, Trebilcock\textsuperscript{48} indicated that it was especially because of these other features that the strict liability regimes in the United States were experienced as “crushing”. Hence, the final judgement on the efficiency of strict liability may also depend upon these other issues, such as causation (which is a particularly important issue in the context of medical malpractice) and the magnitude of the damages awarded.

C. Liability as an Incentive System for Medical Malpractice

1. Negligence versus Strict Liability for Medical Malpractice

How do these findings of economic analysis apply to medical malpractice?

As far as the economic effects of liability are concerned, we already mentioned that economists consider the liability system as one which should give incentives for appropriate care to health care providers\textsuperscript{49}. The idea that the potential injurer will have to compensate his victim is supposed to have a preventive effect\textsuperscript{50}. This preventive effect of liability rules is also stressed in the context of health care. Thus, medical malpractice can basically be considered as one possible regime to give appropriate incentives to health care providers\textsuperscript{51}.

\textsuperscript{49} These basic thoughts have been developed among others by S. Shavell, Theoretical issues in medical malpractice, in: S. Rottenberg (ed.), The economics of medical malpractice (1978), 35–64.
\textsuperscript{50} Compare G. Schiemann, Argumente und Prinzipien bei der Forbildung des Schadensrechts (1981), 185.
\textsuperscript{51} This preventive function of liability rules is, in the context of health care, among others stressed by H. Kozioł, Die Arzthaftung im geltenden und künftigen Recht, in: W. Raderm (ed.), Haftungsrechtliche Perspektiven der ärztlichen Behandlung (1997), 22 and also H. Kötz, Deliktsrecht (6th edn. 1994), 46–52.
32 One of the questions, which we already addressed, is whether a negligence rule or a strict liability rule should be used to provide these incentives for optimal care to the health care provider. Economic analysis has shown that in principle both a negligence rule and a strict liability rule can give these appropriate incentives for optimal care. If strict liability is considered as a legal rule which forces the injurer to compensate the damage of the victim irrespective of the care he took, the injurer will invest in care to the point where the marginal costs of preventive measures equal the marginal benefit in a reduction of the potential loss. Hence, strict liability reaches an efficient solution. However, strict liability assumes that the injurer (in this case the health care provider) knows the risks and the optimal preventive measures to reduce the risk.

33 We already indicated that on the basis of the theoretical literature it is not easy to provide a clear cut answer to the question whether medical malpractice should be governed by a negligence or by a strict liability rule. An argument in favour of strict liability is that the health care providers might have far better information to control the risk than the courts. Courts may indeed often lack the technical knowledge to lay down a standard of optimal care in the provision of the health care services. However, in practice the care exercised is obviously not only governed by liability rules, but also determined through (self) regulation. Hence, these regulatory standards may well guide the courts in applying the negligence standard. Moreover, we equally indicated that strict liability is efficient only if the potential insolvency problem can be cured. Since health care providers could easily cause damage to an amount that may largely exceed their personal assets, a risk of underdeterrence emerges. Thus strict liability can only be considered efficient if it is combined with instruments to take care of the insolvency risk, such as insurance. Moreover, we have to remind ourselves of Trebilcock’s warning that a strict liability regime may well be considered as “crushing” (and perhaps as inefficient) if it is considered with other features such as alleviation of the burden of proving causation. Since causation is effectively often the most important issue in medical malpractice a presumption of causation may well constitute a more important expansion of the scope of liability than the mere shift from negligence to strict liability.

34 As we indicated above, in legal practice “grey areas” may exist between negligence and strict liability. This is for instance the case with a negligence rule with a reversal of the burden of proof. Precisely since for medical malpractice there is probably not a clear cut test to give for negligence or strict liability one may expect that legal practice moves somewhere in this grey area between negligence and strict liability.

52 See A.M. Polinsky (supra fn. 14), 339 and S. Shavell (supra fn. 12), 8.
53 See on this information advantage of strict liability above no. 27.
54 We refer to a reversal of the burden of proving negligence (or compliance with a standard of care), not a reversal of the burden of proving causation.
Obviously there are reasons why in some legal systems judges or legal doctrine plead in favour of either strict liability or a reversal of the burden of proof under negligence. The reasons for these expanding liabilities of health care providers often have to do with the wish to provide full compensation to victims. Economic analysis, however, primarily considers liability rules as instruments of accident prevention, so that victim compensation is not the primary goal of tort law, at least in traditional economic analysis. Second, there is a certain danger, as will be shown below, to over-stress the compensatory functions of strict liability in that a strict liability rule can even lead to a negative redistribution. Third, a strict liability rule also has its weaknesses as a compensation mechanism. Without solvency guarantees (e.g. compulsory insurance) a strict liability rule does not necessarily provide a guarantee of optimal victim compensation. In addition, even strict liability does not guarantee an automatic liability for the injuries since the victim will still have to prove causation which, as we indicated, is a crucial point in medical malpractice. Moreover, in case of insolvency, strict liability may lead to under-deterrence. And another important question one has to ask in the context of medical malpractice is who ultimately bears the costs of an increased level of compensation to the victims. This is an issue which is precisely addressed in the literature following the Coase Theorem, which will be discussed in the next section.

2. Efficient Care in Medical Malpractice

Before doing so it is important to discuss how a standard of due care would have to be applied in the context of medical malpractice if a negligence rule were adopted. In this respect we have to refer to the discussion above, indicating that the legal system will use a generalised abstract due care level, a *bonus pater familias* standard, whereas a differentiation of the standards of care may be efficient from an economic point of view.

The differentiation of the standard of care in case of professional malpractice fits easily into the economic model of tort law. The costs of classification of this case are relatively low. It is indeed easy for the judge to establish whether a certain defendant belonged to a specific profession. Since the costs of classification are low, the judge can easily apply the higher standard of care that corresponds to the professional activity. In other words, the benefits of classification (the ability of requiring a higher level of care from the professional) in this case largely outweigh the marginal costs of classification. This classification in several standards of care in case of professional liability corresponds with the economic model. It indeed forces the professionals to follow the standard of care that corresponds to their superior knowledge. Hence, case law can require the efficient standard of care for the specific professionals.

55 See supra no. 17.
D. The Meaning of the Coase Theorem for Medical Liability

1. Coase Theorem: Introduction

Viewing the above explained advantages and disadvantages of strict liability, one may question the effectiveness of expanding (through e.g. strict liability) liability (to a “better” protection of victims) for health care providers. Economic analysis points to the fact that in many cases a patient is bound with the health care provider via a contract. In that case Nobel Prize winner Ronald Coase taught that an efficient allocation of resources will always follow, irrespective of the contents of the legal rule, as long as the transaction costs are zero. The basic idea of the so-called “Coase theorem” is that, when parties are fully informed, the liability rule will have no effect on the preventive measures to be taken. A classic application of the Coase theorem can be found in the context of product liability (for instance, for pharmaceutical products). In that case the producer and the purchaser (assuming that this is the potential victim) are bound to each other via the price-mechanism. This price can provide an indication concerning the division of risk (of harmful effects of the product) between the parties. In this case the Coase theorem teaches that when parties are fully informed, a change in the liability regime will have no effect on the preventive measures. If the producer were not liable for the damage caused by the pharmaceutical product, the well-informed purchaser would nevertheless add the expected loss (the risk) to the market price and hence only buy the product, taking into account the full price (which includes the risk) of the product.

The consequence of the Coase theorem therefore is that when parties are well informed, the contents of the liability rule do not affect the level of prevention. As long as the conditions for the Coase theorem are met, both an efficient output and efficient care will follow, irrespective of the legal rule applied (no liability, negligence or strict liability). Important too is that under these conditions the choice of the legal rule also has no distribution consequences. Indeed, in a Coasean setting choosing a strict liability rule e.g. to “protect” the consumer seems meaningless, since the expected accident costs will then still be passed on to the consumer through the price-system. Therefore, in these circumstances, legal intervention does not seem very useful. On the contrary, as will be shown below: if consumer groups are not homogeneous, but the expected accident costs differ for every consumer, the introduction of a strict

56 This is the case even though medical malpractice is not always regulated on the basis of contractual liability.
57 R. Coase, [1960] JLE, 1–44.
58 This has been developed by W.Y. Oi, The Economics of product safety, [1973] Bell Journal of Economics (BJE), 3–28.
60 A common example is that the potential loss caused by the burning of one’s hand through a defective toaster will be much higher for a high quality piano-player than for the average economist.
liability rule might have adverse effects, since safer products might disappear from the market\textsuperscript{61}. The reason is that the producer does not know which are the consumers with high-expected accident costs. He will therefore set the quality (care) standard at the level of an average consumer and safer products will disappear from the market\textsuperscript{62}. These problems caused by heterogeneous consumer groups, which will be explained in more detail below, have often been advanced as an argument against strict liability.

2. Strict Liability for “Full Internalization”

The papers using the Coase theorem to analyse product liability were subjected to criticism, since it was held that the theorem does not help since the conditions will almost never be fulfilled in a product liability setting. One argument often heard in this respect is that the consumer has no information with respect to the accident risk\textsuperscript{63}. Although consumers can indeed be poorly informed with respect to some risks\textsuperscript{64}, this is certainly not true for all risks. The consumer might be well aware of the risks caused by the use of e.g. a lawn mower. In addition, it should be mentioned that often information could be provided through legislative measures. The whole regulation concerning the duty to disclose risks concerning pharmaceuticals obviously aims at overcoming this information deficiency. If this were successful, the conditions for the Coase theorem could again be met. Only if a consumer underestimates the accident risk and the information-deficiency cannot be overcome through regulation or otherwise, should a liability rule induce the producer or health care provider to spend enough care on the safety of his products or services. If consumers remain uninformed, non-liability will certainly not lead to an efficient result since producers will have no incentive to invest in safety equipment if this benefit (reduction of expected accident costs) is not recognised by the consumer. As we indicated in the previous section, it is generally held that a fault-rule can induce the producer in such case to take efficient care, but only a strict liability rule will also lead to an optimal output of products or an optimal activity level in case of services. In this respect strict liability is considered to be a good remedy for underestimation of the risk by consumers\textsuperscript{65}.

\textsuperscript{61} This point has been made by W. Oi, [1973] BJE, and has been developed by M. Adams, Wohltat oder Pflege – eine Ökonomische Analyse, [1987] Betriebsberater (BB), annex 20 to journal 31, 5–10.


\textsuperscript{64} E.g. the carcinogenic effects of asbestos or the particular side effects of certain drugs.

Up to this point we have reconfirmed that the economic theory does not jump to hasty general conclusions with respect to the desirability of a fault or a strict liability rule. Such answer depends on whether the victim is a consumer or a third party and if the former, a further distinction is made depending on whether or not the consumer underestimates the risk of a product-accident. However, if consumers do underestimate the risk or if the harm is caused to third parties, the literature clearly points to the advantages of strict liability from an efficiency viewpoint. These are mostly referred to as the full internalisation of the harm through the strict liability rule. However, the economic theory does not stop here. Until now it was assumed that the behaviour of the victim had no influence on the accident risk, that both producers and consumers were risk-neutral and that the market was perfectly competitive. Other publications have shown that when these assumptions are relaxed the advantage of strict liability, expressed earlier, does not prevail any longer. The Coase theorem has a number of interesting consequences, which are more particularly related to the distributive effects of the choice for a particular liability regime.

3. No Distributive Advantage of Strict Liability

Many lawyers favour the introduction of strict liability since it would protect consumers and patients. It is, in the context of product liability, even argued that strict liability would be necessary "to restore the broken balance between producers and consumers". The Coase theorem, however, teaches that also from a distributive point of view there will be no difference between a negligence rule and a strict liability rule, as long as there is a contractual relationship between the potential injurers and the potential victim. Fully informed consumers will only take into account the full price of the product (cost price + expected accident costs) and will base their purchase decision upon this full price. Even if the legislator would like to protect the consumer by introducing a strict liability rule, the producer will still add the expected accident costs to the cost price. The consumer will again pay the full price since the expected accident costs are passed on to him, which is reflected in a higher market price. Since producers and consumers are bound through the price mechanism, every shift of liability to the producer will be passed on to the consumer. The same applies in the case of the purchase of a service (health care) by the patient. In this setting a legal intervention to redistribute wealth to the consumer by introducing a strict liability rule seems therefore useless.

When applied to a situation in the medical practice, the result of the Coase theorem is that when both parties are well-informed, the question if the producer is liable or not, has no influence on the carefulness to be expected from him. It is clear that the presumptions of the Coase theorem in practice, especially with medicine, as we will discuss below, are often not met. The basis for this conclusion is simple: Coase assumes that the consumer (= patient) is in-

formed concerning the health risks, but this information is often not provided for. This lack of information is precisely the reason for stringent regulations concerning the quality of medicine and that those who may sell medicine (the pharmacist) need to have a license to do so. These regulations aiming at information provision and improved quality of products and services somewhat remedy the lack of information towards patients. Nevertheless the fact remains that in this example the purchaser (and presumed user) of the medicine or health care service and the seller (producer or pharmacist) are, from an economic perspective, in a contractual relation with each other. This has important consequences. Often it is argued that the consumer will be protected through strict liability. The application of the Coase theorem in this case shows that a protection like this will only work to a certain extent. The producer or pharmacist will always have the possibility, through extended protection e.g., to charge the extra expense to the consumer in the price of the product. This finally means that an increased consumer protection has the direct effect that the price of the product bought from the seller, for example medicine, will increase, whereby the consumer himself in the end pays for the offered protection. The consequence of the fact that in this case injurer and victim stand in a relation because of the price-mechanism, is that it will be difficult to effect a redistribution to the advantage of the victim, simply because the extended protection is actually paid by the consumer himself through the extended price he is paying in the end. This result notably holds in the area of liability for pharmaceuticals, but also in any case where a provider can succeed in passing on the costs of an increased “protection” via the price mechanism.

4. Potentially Negative Redistributional Effects

So far it has been presumed that the group consumer is homogeneous, meaning that all consumers suffered a comparable damage. This is not always the case in reality. Consumer groups are heterogeneous. A major disadvantage of extended producer liability (through strict liability) is that it will lead to relatively high (medicine) prices because of this extended protection, whereas this will only help a limited group. The heterogeneity within the group of consumers is caused by the differences in incomes by the victims. When consumer groups are heterogeneous, the introduction of strict liability will even have adverse distributional effects. Indeed, the most important part of damages is lost income. The expected damages are therefore of course higher for high-income consumers than for low-income consumers. The producer will, however, take into account average expected damages and will add these to his cost price to get one single market price for all consumers. The effect is that when low-income consumers still buy the product, they “pay” for the expected damages of the high-income consumers. Therefore strict liability in a product liability setting creates a redistribution from poor to rich consumers. When we proceed from this point of view and assume that for example a defective medicine has

been delivered, causing the victim to be disabled to work, then it will be clear that the damage suffered by a “poor” victim (that is a victim with a small income) will be much smaller compared to a “rich” victim with a higher salary. Still all victims will be fully compensated because of the system of strict liability, with the difference that the protection will serve the richer more than the poor. The perverse consequence is that an extended consumer protection in this case will thus lead to increased prices of medicine for all consumers, whereby this extended protection will work more in favour of the richer victims.

This result proves once again one of the important issues stressed in the economic literature, being that strict liability even leads to a negative distribution. This corresponds with the American literature concerning the so-called insurance-crisis in the United States of America. For the same reason Priest argued that the product liability explosion in the US has especially hurt the low-income groups⁶⁸, because the increase in the premium mainly has been caused by the higher income-groups. This shows that we have to be very careful in assuming that strict liability would be beneficial for “consumer protection”.

5. Extended Protection for Third Parties

It can be pointed out that traditional product liability law only applied to those who bought a service or product, because of the privity of contract only “consumers” (contract parties) could be compensated if damage occurred. The European Product Liability Directive of 1985 extended this liability also to the advantage of third parties affected by defective products bought by other parties.

A detail which should also be mentioned is that the notion “consumer protection”, which is so often used in the considerations preceding the Directive, is in fact misleading. A “consumer” is a purchaser, who buys a product. The Directive is, of course, not at all limited to a protection of the buyer of the product. Also third parties that suffer harm from a defective product can benefit from the protection of the Directive. Indeed, for many Member States the main improvement of the Directive is that also third parties now benefit from the strict liability of the producer. The damage suffered by someone who stood in a contractual relationship with the producer (the “consumer”), was often already subjected to strict liability of the producer. In addition, the damage caused to a victim “as consumer” being the damage caused to the purchased product itself, is excluded by the Directive⁶⁹. For these reasons it appears that the term “consumer protection”, which is so often used with respect to the Directive, is not that well chosen. What is meant is the protection of the victim, whether this is a consumer or not.

a) Some Economic Effects of the EC Directive

Apparently the drafters of the EC Directive judged that the best way to realise the "protection of the consumer", of which they spoke so much, was to implement a generalised strict liability rule for harm caused by a product defect. The first thing that is striking is that the Directive advances one single rule to deal with all product accidents, whereas the economic theory, as we have seen in previous sections of this report, is much more balanced and detailed. Depending upon whether or not the victim is a consumer, whether the latter has information on the product risk, what the influence of insurance is etc. economic theory would only hold that a strict liability rule is efficient in some cases, but certainly not in all.

At the policy level it is, using all the relevant criteria, almost impossible to give advice to the legislature to use one single rule for all product accidents. There is even a question as to whether the legislator would have to define such a rule. An obvious alternative would be to let the judge decide what type of liability rule should govern the product accident. Thus case law could take into account all the relevant criteria and the application of the strict liability rule could be limited to certain product accidents. The Directive, which introduces one strict liability rule for all product accidents, disregards all the various factual situations, which make different liability rules efficient. Thus, by using one single rule inefficiencies will unavoidably be created since the strict liability rule will also be applied where this would be inefficient taking into account the above-mentioned criteria. It is also doubtful whether the savings in administrative costs by using one legal rule outweigh the inefficiencies. The costs of applying the economic criteria for strict liability should not be that high either. Especially if they can easily be recognised (e.g. whether the victim was a consumer or a third party) an individualised product liability system where strict liability will only be applied in some cases could be used by the judge at relatively low costs.

One of the criteria was also applied for a long time in legal practice. Indeed, in most legal systems product accidents, whereby the victim stood in a contractual relationship with the producer, were treated differently from those where the victim was a complete stranger to the producer. One of the reasons for introducing the generalised strict liability in the Directive was that this different treatment was considered "unjust". However, this distinction is quite sound from an economic point of view. As was mentioned above, if the victim stands in a contractual relationship with the producer a Coasean solution is in principle possible if the consumer is well-informed of the accident risk. Such solution is of course always excluded in the event that the victim is a third party.

because of the prohibitive transaction costs. So there is a good economic reason for treating both situations differently.

Of course the EC Directive does not introduce a general strict liability rule for all damage caused by a product. The producer will only be liable when the damage was caused through a defect in the product. It looks like this is more a fault liability since a product is often defective because the producer did not take efficient care. However, liability is indeed strict since the producer will always incur liability when harm was caused by a defect in his product even if he can prove to have taken due care. The producer only escapes liability when the harm was not caused by a defect of the product. Given the broad defect notion in the Directive a product is almost already considered defective by the mere fact of having caused the harm\textsuperscript{71}. Therefore, there is not, from an economic point of view, a substantial difference between strict liability for all damage caused by a product and strict liability for damage caused by a defect in the product as in the Directive\textsuperscript{72}.

It was mentioned above that the general strict liability rule for all activities, which is used in the Directive, might create inefficiencies since it will also be applied in cases where a fault rule would be preferable from an efficiency viewpoint. It should not be forgotten, however, that the purpose of the Directive was the harmonisation of product liability law in Europe. If this should succeed, it could create benefits, which could easily outweigh the disadvantages of the generalised strict liability rule discussed above. Indeed, if a single product liability rule were to be created in all Member States, this could contribute to the creation of equal marketing conditions and thus to the realisation of the internal market. This could bring substantial savings\textsuperscript{73}, which could easily outweigh the inefficiency of the use of a single rule. This, however, will not be the case in practice since it was shown in the previous subsection that the Directive can never bring a harmonisation of product liability law in Europe. So the inefficiencies remain without any compensating benefit for the realisation of the internal market.

b) Distributional Effects

One could go one step further and argue that the Directive does not only create several inefficiencies, but that it is also problematic from a distributive point of view. Many lawyers favour the introduction of strict liability since it would

\textsuperscript{71} One could argue that as soon as a product causes serious injuries it does not provide the safety which a person is entitled to expect and is therefore defective according to article 6 of the Directive.


\textsuperscript{73} Compare R. Van den Bergh, Competition law and competition in Europe after 1992. Will free riders become the heroes of the internal market?, unpublished manuscript.
Economic Analysis

protect the consumer. It is even argued that strict liability would be necessary “to restore the broken balance between producers and consumers”. The Coase theorem teaches that also from a distributive point of view there will be no difference between a fault rule and a strict liability rule. Fully informed consumers will only take into account the full price of the product (cost price + expected accident costs) and will base their purchase decision upon this full price. Even if the legislator would like to protect the consumer by introducing a strict liability rule, the producer will still add the expected accident costs to the cost price. The consumer will again pay the full price since the expected accident costs are passed on to him, which is reflected in a higher market price. Since producers and consumers are bound through the price mechanism, every shift of liability to the producer will be passed on to the consumer. In this setting a legal intervention to redistribute wealth to the consumer by introducing a strict liability rule seems therefore useless.\(^\text{76}\)

When consumer groups are heterogeneous, the introduction of strict liability will even have adverse distributional effects. Indeed, the most important part of damages is lost income. The expected damages are therefore of course higher for high-income consumers than for low-income consumers. The producer will, however, take into account an average expected damage and will add this to his cost price to get one single market price for all consumers. The effect is that when low-income consumers still buy the product, they “pay” for the expected damages of the high-income consumers. Therefore strict liability in a product liability setting creates redistribution from poor to rich consumers.\(^\text{75}\). For the same reason Priest argued that the product liability explosion in the US hurts especially the low-income groups.\(^\text{76}\)

This effect will be stronger still when the strict liability rule does not only apply to harm suffered by consumers, but is extended to damage caused to third parties, as the EC Directive does. In that case the market price will again be higher since the producer will also have to pay for damage caused to third parties. The increased expected damages will again be passed on to the consumer who pays a higher market price for the damage a product he bought can cause to third parties. The consumer has no possibility to pass on this increased price and therefore he, in effect, pays for the protection of third parties.

So the expansion of strict liability to damage caused to third parties redistributes wealth from consumers (of the product) to third parties. Consumers will indeed have to pay for the protection of third parties. This again increases the adverse redistributive effect of poor consumers paying higher prices for the expected damages of third parties. Since lost income is an important part of the damage, again the high-income groups do indeed benefit from the redistribution.

This shows that the generalised strict liability, which has been introduced by the EC Directive, will not only create inefficiencies, but is also based on wrong ideas concerning the protection of consumers.

c) Curing Externalities?

So far, in our analysis of the European Product Liability Directive, we argued that the Directive can hardly lead to a harmonisation of marketing conditions or to a lowering of transaction costs (given all uncertainties and information problems). Another point, made by Van den Bergh, is that the Product Liability Directive can also not be considered to cure the risk of externalising damage caused by defective products since the Directive does not address the issue where such a risk of externalisation may occur. First of all Van den Bergh refers to the fact that such an externalisation may occur in case non-pecuniary losses are insufficiently compensated under national law, which may lead to underdeterrence. Another problem may arise if full scientific proof of a causal link with a certain product defect is required and e.g. proportionate liability is not accepted. The Directive does not cure these problems, since it does not touch upon the issue of causation, nor upon the compensation of non-pecuniary loss, which is explicitly left to the national Member States. Moreover, the Directive allows a financial limit on compensation, which is generally considered inefficient, especially in case of strict liability. Therefore the Directive in fact increases the risk of cross border externalities, so Van den Bergh argues. Finally there is a problem in the Directive as, as we indicated above, it does not apply to retailers. Above it was indicated that a state cannot attract industry with lenient product liability legislation for manufacturers, but it could attract retailers by doing so. The EC Directive does not cure that risk since it does in principle not apply to retailers.

6. Coase Theorem and Its Relevance for Medical Malpractice

At first sight, this Coase theorem discussed in this section did not seem to have much importance for medical malpractice. Indeed, the Coase theorem assumes zero transaction costs, which in the context of medical care means that it is assumed that the patient is fully informed about the risks involved in a treatment. This assumption is often unrealistic. As such, the fact that many legal systems force health care providers to provide adequate information on risks fits into the economic analysis. It may allow patients to make a better-informed choice. The fact that patients often have little information on risks (or are not able to evaluate the risks, even when properly informed) is one of the

82 See M. Faure (supra fn. 44), 374.
reasons why regulation should intervene (in addition to liability) to regulate e.g. the quality of pharmaceutical products.

However, the fact that there is a contractual relationship between a health care provider and a patient (or between a producer and a consumer of pharmaceutical products) is still important from an economic perspective, even in the cases when the patient cannot be adequately informed. Indeed, it is often argued that the consumer (or patient) should be protected via a strict liability rule (or via an extended negligence rule, such as e.g. the reversal of the burden of proof). The Coase theorem teaches that such a protection will often have just a limited impact, at least when the health care provider will have the possibility to pass on this increased patient protection to the consumer via the price mechanism. This means that if e.g. case law or the legislator decide to shift to a strict liability rule, for instance for damage caused through pharmaceutical products, a producer will simply increase the price of his products with the fact that he now bears the potential loss. Thus the consumer in fact pays for the increased protection which is awarded to him.

In sum, since parties are bound to each other via the price mechanism it is effectively difficult to "protect" a consumer/patient via expanded liability rules since this increased protection can be passed on via the price mechanism, so that in the end the consumer pays for it himself.  

The importance of the Coase theorem in the context of medical malpractice is that judges or legislators can, from a legal perspective, try to award an increased protection to victims (e.g. through a reversal of the burden of proof), but that as long as victims and injurers are bound via the price mechanism, the health care provider will be able to pass on this increased protection to the patient via the price of the services. Obviously one should realise that the scope for truly free negotiations between a health care provider and a patient concerning the price of the services and the related risk will in practice not be large. One problem is the lack of information on risks of the patient; another problem is that free negotiations on price are excluded in many health care systems where the prices of health care services have been regulated through the influence of social security. It is nevertheless interesting to point at revolutionary proposals of e.g. Chicago Professor Richard Epstein, who suggested solving the American medical malpractice crisis on the basis of contractual agreements between patients and physicians à la Coase. The basic idea of Epstein is that patients (or their representatives), physicians and hospitals would

---

60 These effects of the Coase theorem have been developed by K. Hamada, [1976] AER, 228–234; see also G. Calabresi, Transaction costs, resource location and liability rules: a comment, [1986] JLE, 67–73. The same point has been made for Germany in an excellent paper by M. Adams, [1987] BB, 1–24.


62 Which is, as will be argued below, a good argument for a regulation forcing health care providers to disclose information on risks.
negotiate *ex ante* on the amount of care (preventive measures) to be taken by the health care provider and on the corresponding division of risks and would agree accordingly on the price to be paid for the services\(^6\).

Although these ideas may not be immediately practicable to solve medical malpractice in Europe, some of the underlying concepts are highly interesting; particularly the fact that an increased protection of patients (via reversal of the burden of proof or a shift to strict liability) has financial effects which will somehow be passed on. The only question is indeed whether the legal regime gives possibilities for passing on these increased costs. In systems where health care providers are free to set prices for their services, an increased protection will lead to increased insurance premiums which will be passed on to the patients. In systems where prices are regulated the increased liability will either lead to higher costs for the social security system (e.g. when public hospitals are made liable) and will thus be spread to the general taxpayers or to increased exposure of insurers for liabilities of their clients. This obviously poses the question what the effects of increased liabilities will be on the insurance level.

E. Regulation

1. Introduction

So far we have examined how medical malpractice could be avoided by using liability rules. The assumption underlying the analysis in the preceding section was that liability rules would give incentives to follow the optimal care to avoid damage to patients. Moreover, we discussed the importance of the Coase theorem pointing at the fact that if it is possible to pass on increased liabilities via the price-mechanism, it will in fact be the patients who, as consumers of the health care services, pay for the increased protection which is awarded to them. In practice, however, it seems clear that liability rules are definitely not the only legal rules that affect the preventive behaviour of the health care provider. In practice the standard of appropriate care in health care will often be defined through regulation. This can either be public regulation, issued by public authorities (this will often be the case with pharmaceuticals, which are subjected to prior approval) or private regulation, issued by professional associations.

---

In this section we will turn to regulation as a device to promote appropriate care of health care providers. We will address the principle question being why in specific circumstances regulation would be a better instrument than tort law to control the care exercised by the health care provider.

Let us examine under what kind of circumstances liability rules may not suffice to deter medical malpractice and a regulatory intervention may be necessary. The choice between regulation and liability rules was thoroughly examined by Steven Shavell in 1984, in a paper in which he advances several criteria that influence the choice between safety regulation and liability rules\(^87\). In the next subsection we will apply these criteria to medical malpractice.

2. Criteria for Safety Regulation\(^88\)

a) Information Asymmetry as a Criterion for Regulatory Intervention

Information deficiencies have often been advanced as the justification for government intervention through regulation\(^89\). For the proper operation of a liability system, information on e.g. the existing legal rules, the accident risk, and efficient measures to prevent accidents (if possible), is a precondition for an efficient deterrence. According to Shavell, the injurer in an accident setting generally has much better information about the accident risk than that possessed by the regulatory body\(^90\). The injurer himself has, in principle, the best information on the costs and benefits of the activity that he undertakes and of the optimal way to prevent accidents. This “assumption of information” will, however, be reversed if it becomes clear that some risks are not readily appreciated by the injurer in an accident setting. This may more particularly be the case if the costs for examining the optimal safety devices would be high. \textit{Ex ante} regulation would thus pass on information to the market (for medicine) or medical practice (for health care services) on optimal safety devices\(^91\) or care.


\(^88\) These criteria are also discussed, but then in the context of environmental harm in M. Faure/D. Grisard, Financial assurance issues of environmental liability, in: M. Faure (ed.), \textit{Deterrence, insurability, and compensation in environmental liability. Future developments in the European Union} (2003), 40–42.


\(^90\) S. Shavell, [1984] JLS, 359.

\(^91\) S. Shavell, [1984] JLS, 359–360; See also F. Silva/A. Cavaleri (supra fn. 65), 297–298.
b) Insolvency Risk

68 If the potential damages can be so high that they will exceed the wealth of the individual injurer, liability rules will not provide optimal incentives. The reason is that the costs of care are directly related to the magnitude of the expected damages. If the expected damages are much greater than the individual wealth of the injurer, the injurer will only consider the accident as having a magnitude equal to his wealth. He will take, therefore, only the care necessary to avoid an accident equal to his wealth, which can be lower than the care required to avoid the total accident risk. This is a simple application of the principle that the deterrent effect of tort liability only works if the injurer has assets to pay for the damage he causes. If an injurer is protected against such liability, a problem of underdeterrence arises. When it concerns a manufacturer an additional argument needs to be considered. Although many manufacturers may have large financial capacities, most of them are structured as corporations, which benefit from limited liability.

69 Safety regulation can overcome this problem of underdeterrence caused by insolvency. In that case the efficient care will be determined ex ante by regulation and will be affected by enforcement instruments that induce the potential injurer to comply with the regulatory standard, irrespective of his wealth.

70 It has, however, been pointed out in the literature that this problem of insolvency causes especially underdeterrence in case of strict liability and less so in case of the use of a negligence rule. Indeed, under strict liability a problem of underdeterrence already arises as soon as the magnitude of the loss will be larger than the injurer’s individual wealth. He will, in that case, consider the accident as one with a maximum magnitude of his individual wealth and therefore only take the care necessary to avoid that accident. That care may be lower than the optimal care. Negligence has the advantage that the injurer continues taking care as long as his costs of care are lower than his wealth, since taking care will free the injurer (health care provider or manufacturer in case of medicine) from liability. This, therefore, once more balances the conclusion that strict liability may be the efficient rule for (product) liability. This is true on the condition that the insolvency problem can be cured, otherwise a risk of underdeterrence might arise.

71 Moreover, this insolvency problem may generally be an argument in favour of ex ante regulation by government whereby a safety standard is set ex ante and enforced with sanctions. If, however, insolvency is the reason to introduce reg-

---

93 S. Shavell, [1986] IRLE, 43–58. Above we mentioned that insolvency causes especially a problem under a strict liability rule, but less so under negligence.
94 If insurance would come into the picture it could overcome the problems of underdeterrence, provided that the moral hazard problem, caused by insurance, can be cured.
95 This point has been made by W. Landes/R. Posner, Tort Law as a Regulatory Regime for Catastrophic Personal Injuries, [1984] JLS, 421–422 and by F. Silva/A. Cavaliere (supra fn. 65), 296.
ulation, these sanctions should obviously be preferably of a non-monetary nature.

c) The Threat of a Liability Suit

There may, however, be a third argument which is probably the most powerful in favour of safety regulation in case of accident, being the risk that injurers may, for many reasons, escape a liability suit. Some activities can cause considerable damage, but even so a lawsuit to recover these damages may never be brought. If this were the case, there would of course be no deterrent effect of liability rules. Therefore, the absence of a liability suit would again be an argument to enforce the duty of efficient care by means of safety regulations rather than through liability rules. The reasons why a lawsuit is not brought, even though considerable damages have been caused can be manifold.

Victims may have difficulties proving that their loss was caused by a product defect; this may be due – *inter alia* – to problems of proving causation. The most important problem is probably latency, being the fact that a long time lapse may happen between the moment that the medicine was put on the market, and the moment that the damage occurs. If e.g. harm only occurs 20 years after a product was brought on the market, the manufacturer may already be out of business and can hence not be reached by tort law any longer.

A related problem is that it is often hard to prove that a causal link exists between the medical treatment and the illness that occurred. The burden of proof of a causal relationship becomes more difficult with the increasing passage of time since the medical care. Often a victim will not recognise that the harm had been caused by a tort, but might think that his particular ailment, e.g. cancer, had a "natural cause", associated with general ill health.

For all these reasons a liability suit might not be brought and hence safety regulation is necessary to ensure that the potential injurer takes efficient care.

d) Administrative Costs

When examining the pros and cons of liability versus regulation, the administrative costs of both systems should also be compared. Liability rules are clearly costly in terms of time for both parties and in court fees. A part of these costs is borne by the whole community, such as e.g. the cost of the legal system, fees for the judge etc. Regulation produces costs for the community,

---

including the costs of making the regulation, setting the standards, passing the statutes etc. and of subsequent enforcement\textsuperscript{101}.

77 In this respect the liability system seems to have an advantage: the administrative costs of the court system are only incurred if medical malpractice has actually happened. The main advantage of the tort system is that a lot of accidents will be prevented by the deterrent effect of being held liable and having to pay damages to the victim. In case of safety regulation the costs of passing the regulation and of enforcing it are always there, whether there are accidents or not.

3. The Need to Regulate Medical Malpractice

78 After having discussed these criteria for regulation\textsuperscript{102}, I will now discuss the question of how these criteria relate to medical malpractice. If one takes the criteria for safety regulation discussed above and applies them to the potential risk caused by medical malpractice, there is no doubt that liability rules alone are not sufficient.

79 If one looks at the first criterion, that of information, it must be stressed that an assessment of the risks of a certain activity (e.g. medicine producers) often requires expert knowledge and judgement. Small pharmaceutical companies might lack the incentive or resources to invest in research to find out what the optimal care level would be. Also, there would be little incentive to carry out intensive research if the results were automatically available to competitors in the market: this is the well-known "free rider" problem. This problem can partially be countered by legal instruments granting an intellectual property to the results of the research. However, the problem remains that it may not be possible for small companies to undertake studies on the optimal safety standard of medicine for preventing major side-effects. Therefore, it is absolutely necessary to set the optimal safety standard by governmental regulation. The same point can be made with respect to the care exercised by a physician. The costs to find out what the appropriate standard of medical care is may be quite high for the individual doctor and certainly for a judge who would have to examine this in the context of a tort suit based on negligence. There may be economies of scale advantages if the appropriate standard (also as far as equipment is concerned) were set after research by a regulator or a professional body. In that case the costs are incurred only once and can be passed on to all health care providers in the market via regulation.

80 Also, the insolvency argument points in the direction of regulation. Medical product defects can be the responsibility of pharmaceutical firms that have assets that are lower than the amount of damages they can cause by a defective product. Moreover, most firms have been incorporated as a legal entity and

\textsuperscript{101} S. Shavell, [1984] JLS, 363–364.

\textsuperscript{102} These are often referred to as 'public interest' criteria for regulation to contrast them with 'private interest' explanations for regulation, as advanced by public choice scholars.
therefore benefit from limited liability. Hence, the individual shareholders are not liable to the extent of their personal assets, but a creditor of the firm can only lay claim to part of all of the total assets purchased in the firm by the shareholders. The same is true for health care providers; they can potentially cause losses with a magnitude that can easily be higher than their personal assets. Hence, there is a real insolvency risk.

Also the chances of a liability suit being brought for damage caused by medical malpractice is naturally very low. A problem is that it is often hard to prove that a causal link exists between an activity and a type of damage, being the medical treatment and the illness that occurred. The burden of proof of a causal relationship becomes more difficult with the increasing passage of time since the medical care took place. Often a victim will not recognise that the harm had been caused by a tort, but might think that his particular ailment is simply the result of his general health condition. Therefore many cases of medical malpractice are never recognised as such and never give rise to liability. For all these reasons a liability suit might not be brought and hence safety regulation is necessary to ensure that the potential injurer takes efficient care. The same problems can occur with medical products.

For these reasons it is clear that some form of (government) regulation of medical products and of medical malpractice is necessary. To reformulate: this shows that liability rules alone cannot suffice to prevent harm. Another question will be whether this necessarily implies that medical malpractice should necessarily solely depend upon regulation or whether regulation can still fulfil a supplementary role.

4. Self-Regulation

For the liability rule to work as an instrument of prevention it requires informed parties or judges, as we have seen above. The judge will determine the standard of care to be taken by the physician in case of a negligence rule, and the potential injurer in case of strict liability. An alternative to give incentives for prevention of medical malpractice is not only to trust liability rules but to impose an ex ante set standard of care on physicians or hospitals. This standard can be set through governmental regulation. In that case this normative standard of care needs to be followed at any time by the legal subject under the threat of (administrative or criminal) sanctions.

Another alternative is not to trust governmental regulations, with the argument that the government will not have the appropriate knowledge and information concerning an appropriate medical care to set the optimal standard of care. Due to this fact, the medical profession has set its own optimal standard of

---

103 This insolvency risk is, however, somewhat remedied through liability insurance. That may, however only constitute a partial remedy, since insurance does usually not provide unlimited coverage.

care through a so-called self-regulation. In most countries the care to be followed by a physician is primarily determined through disciplinary rules, set by the professional organisation. Non-compliance with the standard set means that disciplinary sanctions will be imposed on the individual doctor(s). In some legal systems self-regulation and government regulation are combined and in some cases the non respect of disciplinary rules set by the profession is even sanctioned with public means.

These self-regulation rules do of course have the advantage that they are set by those who have the best information on appropriate care (the physicians themselves), but they can obviously lead to a lot of problems as well. Especially the so-called moral hazard creates a problem, which means that it is not naturally guaranteed that the medical profession itself has sufficient motivation to keep the standard high. Economists have indeed pointed at the fact that although some regulation seems necessary to cure the information deficiency with patients, there is always the risk that this regulation at the same time creates inefficiencies e.g. by limiting competition and limiting market entry. That might lead to high prices and, therefore, the contents of these self-regulation rules should always be followed with criticism.

It is important to notice that both liability rules and (self-) regulation can give the physician the incentive to follow the optimal standard of care. There needs to be made a distinction however, that the optimal standard of care set through self-regulation has to followed whether an accident occurs or not. The (administrative, criminal or disciplinary) sanctions can be imposed as soon as the standard of care is violated, even when there is no damage in a concrete case. Liability rules on the other hand become effective ex post, only after an accident has happened.

5. Liability and (Safety) Regulation

Although we just argued that medical care should be subjected to regulation, one can equally argue that rules of regulation alone are not sufficient in preventing damage. First of all the effectiveness of a rule of regulation is determined by its implementation in practice. In the latter case this does not seem to be sufficient. Liability is therefore necessary. Secondly, it needs to be pointed out that (self-) regulation cannot cover all possible situations and needs. Indeed, the effectiveness of regulation is dependent upon enforcement, which may be weak, more specifically in the context of self-regulation. The profession might not always have optimal incentives to highlight bad practices and punish “quacks” adequately. This tendency might to some extent be overcome by combining reg-

---


ulation and liability rules. Indeed, liability rules do also have the advantage that they are enforced by patients. Their incentive for an impartial enforcement might often be better than the incentives of the professional body.

For all these reasons, liability rules should still play an important role in preventing medical malpractice. This obviously raises important questions with respect to the relationship between regulation and liability rules, which will be addressed in this section.  

a) Violation of Regulation and Liability

The first question to be answered in that respect is whether a violation of a regulatory standard of care should automatically be considered a fault under tort law and thus lead to liability of the medical caretaker.

In most legal systems the standard of care to be taken by medical care providers engaged in possibly risky activities is not only determined by the court system within the framework of tort rules, but also by the legislator, and, as we just indicated, also by professional organisations. Economic analysis of law has, as we noticed in the previous section, advanced several criteria to justify the implementation of safety regulation: information problems, insolvency, absence of deterrent effect of tort suits in some cases and administrative costs. Given the existence of regulation and liability rules in combination, the question arises how the influence of legislative safety regulation influences the tort system. In many legal systems every violation of a regulatory standard will automatically result in a finding of “fault”. This means that as soon as there is a violation of the regulatory standard, the blameworthiness requirement is met and the violation stands in a causal relationship with the damage, the injurer has a duty to compensate.

b) Economic Analysis

The question arises whether this per se rule of liability in case of violation of regulatory standards fits into the economic analysis. Shavell argues that noncompliance with a regulatory standard should not automatically result in a finding of negligence. Following the standard might be inefficient for some injurers. The injurers for whom following the regulatory standard would only be possible at high costs, should not be held to follow this standard, since it would create inefficiencies. By making every violation a “fault”, those injurers are given wrong incentives. They have to comply with the regulatory standard, even if it were inefficient for them to do so. But of course

---

108 See in this respect also M. Faure/D. Grimeaud (supra fn. 88), 55–58: where an application to the area of environmental harm is provided.


this "inefficiency" is partially compensated by the considerable administrative costs resulting from a detailed individualisation. The question, therefore, arises whether some injurers should not be held liable if they violate the regulatory standard.

The problem can be compared with the *bonus pater familias* standard\textsuperscript{111}. Although a detailed individualisation of standards of efficient care would be the optimal solution in a first-best world this is often impossible given the costs of an individualised standard setting. Therefore, the legal system sets the required level of care at an average level, the so-called *bonus pater familias* standard. The same can be said for regulation. Atypical physicians might be able to avoid a loss at lower costs, e.g. because they pose lower risks than normal. The regulator cannot identify atypical parties and therefore a single regulatory standard will be used. Hence, it is best for those parties not to comply with the standard. If various groups can be identified at low costs, a separate standard for a certain group is efficient as long as the gains from selecting a further group outweigh the further administrative costs. In most cases, however, the regulator will not have the possibility of identifying atypical parties that might be able to avoid a loss at lower costs, for instance because they pose lower risks than normal. Therefore, a single regulatory standard will be used\textsuperscript{112}. The result is that one could, therefore, argue from an economic point of view that a failure to satisfy the regulatory requirement should not necessarily result in a finding of negligence, so as to avoid some parties, who pose lower risk, taking wasteful precautions. Nevertheless, many legal systems consider a breach of a regulatory duty, such as ethical standards, automatically as fault. This may be due to the information advantage of regulation. Regulation passes on information to the parties about the efficient average standard of care. At the same time regulation also gives information to the judge who has to evaluate the behaviour of the physician *ex post* in a liability case. The judge might often lack the necessary information to find out whether in a particular case the health care provider should not be held to have followed the regulatory standard, for example because he posed a lower risk than usual. This explains why general standards are applied to define negligence\textsuperscript{113}. An exception to this rule will probably only be made if the judge can, at low costs, identify specific groups of injurers and set a separate standard for that group.

One can also understand the application of this rule in the context of medical malpractice, since information deficiencies were an important reason to introduce regulation in the first place. It is understandable that this regulation guides the judge who has to define whether the health care provider took due care in a negligence case. The regulation hence provides information to the judge on the (minimum) care that the health care provider is expected to fol-

\textsuperscript{111} See explanation of 'bonus pater familias' (supra fn. 27) and see supra no. 17–18.

\textsuperscript{112} Compare for the *bonus pater familias*-standard, R. Posner (supra fn. 26), 183–184 and S. Shavell (supra fn. 12), 74.

\textsuperscript{113} See on this issue also S. Rose-Ackerman, *Rethinking the progressive agenda, the reform of the American regulatory state* (1992), 127.
low. If that standard has been breached, the judge will consider such a behaviour as negligent. Hence, regulation guides the judge in a liability trial, which does not seem inefficient. This is even the case if the regulation is not set by a public authority, but by the professional organisation. Indeed, one cannot see what incentive they would have to set the care level inefficiently high. Since these care levels required by the professional bodies will in most cases only be minimum levels of care one can understand why their violation is automatically considered a fault.

c) Justificative Effect of Compliance with Regulation

Above we have explained that to a large extent the prevention of medical harm is also achieved as a result of regulation. Whereas according to tort law in many legal systems a breach of a regulatory standard results automatically in a finding of negligence, the opposite is not true: following a regulatory standard does not exclude a finding of liability. In medical malpractice law this is particularly important, since the criteria for a standard of (minimum) care that a physician has to fulfil are largely regulated. The medical profession could argue that as long as this level of care is met, the physician is not liable for negligence in tort law. If compliance with a regulatory standard automatically resulted in a release from liability, the potential injurer would have no incentive to invest more in care than the regulation asks from him, even if additional care could still reduce the expected accident costs significantly.¹⁴

A first reason to hold an injurer liable (if the other conditions for liability are met), although he has followed the regulatory standard, is that this standard is often merely a minimum. This is most certainly the case with self-regulation. The complete "compliance defence" prevents any incentive to take precautions in excess of the regulated standard.¹⁵ Exposure to liability will give the potential injurer incentives to take all efficient precautions, even if this requires more than just following the standard of care which is laid down in regulations. Second, tort law can also be seen as a "stopgap" for situations not dealt with by regulations.¹⁶ Moreover, liability can function as an important additional deterrent for the cases in which the regulatory standard had been set too low. Also, if following a regulatory standard of care would lead to a complete defence, this would substantially reduce the victims' rights to compensation. This makes clear that the exposure to liability notwithstanding the regulatory standard is an important guarantee that the medical caretaker will take efficient care.

¹⁶ S. Rose-Ackerman (supra fn. 113), 123.
d) No “Second Guessing”

Therefore, following the conditions of the regulatory standard should not have a justificative effect in tort. The opposite may only be true if it were clear that the legislator (or professional bodies) took into account all potential medical harm when setting the regulative standard of care. In such case a judge in a civil liability suit should not be “second guessing” efficient legisliative decisions. It is, however, rare that the legislator (or professional bodies) will be able to take ex ante all these interests and possible damages into account when setting the standard of care conditions. Hence, as a general rule, following regulatory standards should not free one from liability; the opposite would be the exception. This is the case both under negligence as well as under a strict liability rule. Indeed, holding an injurer liable, notwithstanding the fact that he followed regulatory standards will play an important role under a strict liability rule, since this will lead the injurer to take efficient care and adopt an efficient activity level, i.e. to take all efficient measures to reduce the potential accident costs, although this might require more to be done than the regulation requires. Under a negligence rule this case law is also significant if the efficient care standard (which is assumed to be equal to the due care standard required by the legal system) is higher than the regulatory standard.

F. Empirical Evidence

1. Defensive Medicine

Of course one needs to keep in mind that this is just a theoretical approach why liability and regulation are the most efficient solution to give medical care providers optimal incentives to prevent medical malpractice. That does not mean that the current laws in European legal systems are efficient. The self-regulation and the insurance system of public health service can only give reason for a limited incentive. Courts in some cases set the optimal care standard too high, consequently physicians will administer precautionary medical treatments that have minimal expected medical benefit out of fear of legal liability. This problem, called “defensive medicine”, caused problems in the United States of America\textsuperscript{117}. Kessler and McClellan examined in their empirical study the extent to which managed care and liability reforms interact to affect the cost of care and health outcomes\textsuperscript{118}.

2. Safety Regulation in Practice

a) Empirical Analysis

When Shavell’s criteria for safety regulation are applied to medical malpractice, one can easily note that a strong argument can be made that the efficient


\textsuperscript{118} See D.P. Kessler/M.B. McClellan, [2000] Working paper series. The examined group of patients was elderly Medicare beneficiaries with cardiac illness.
care to be taken to avoid medical malpractice should also be fixed *ex ante* by regulation. In many cases this regulation consists of licences (for medicine) or correlative disciplinary sanctions by state medical boards (for medical caretakers) in which an administrative authority fixes a care-taking standard which must be followed by the potential injurer. An improvement of the quality of medical care will mostly be affected by imposing more stringent care-taking standards in safety measures (or administrative licences for producers and providers of medicine). Hence, the general requirement that the medical care level is controlled through safety measures (and licences) and that the quality and quantity of the medical care is regulated by the conditions of the safety measures (and in licences), is a cornerstone of medical law.

Although the standard of care for health care services is basically controlled through these (self-) regulatory standards, in individual cases there can still be damage. Then again liability under tort law comes into the picture and the question is raised of the influence of regulation on the liability system and *vice versa*.

b) Output Analysis

Based on the findings of a number of studies focused on the change in health care delivery attributable to the treatment of medical liability and the impact on medical malpractice also seen from the perspective of regulation, Dewees, Duff and Trebilcock made an output analysis in their book “Exploring the Domain of Accident Law”.

Summarised, they examined different studies to show some impact of civil liability on medical practice, which was shown to some extent, but they found it impossible to reach adequate conclusions because there was inadequate evidence on both the costs and benefits of liability-induced changes in medical malpractice, and in the precise character of those changes themselves. But the limited evidence, according to Dewees, Duff and Trebilcock, suggests two areas in which the civil liability system may have had a noticeable impact and where the benefits of these practice changes probably outweigh the costs: physician discussion of treatment risks and alternatives with patients (as well as increased record keeping), and institutional programmes for injury prevention (e.g. risk management programmes).

c) Individual Competence

Dewees argues that the primary method of ensuring competence in the medical field is through licensing. This is the well-known issue of self-regulation, discussed above. In America a principal quality-control assurance mechanism has always been the regulation of new entrants into the medical

---


120 See D. Dewees/D. Duff/M. Trebilcock (supra fn. 119), 112.
field by medical boards. Three problems are associated with quality-control assurance mechanism of this kind of regulation:\(^{12}\):

Devising education and training requirements that can demonstrate, by examining the new entry's academic performance, an appropriate level to practice in the medical field, as it is difficult to determine from student academic performance if someone is going to be a successful doctor.

Due to rapid changes in knowledge and technology within the medical field, professional obsolescence is a real concern with physicians. To maintain the required level of competence and quality of care there are two strategies proposed - mandatory periodic re-evaluation and ongoing medical education as a condition of re-licensure.

There is evidence that professional control over licensing is used to achieve economic advantages for members of the profession (e.g. Elton Ray-ack, 1983). One study concludes that when market conditions worsen, licensing boards tend to lower the number of licenses awarded to new entrants to keep the position of the licensed competitive on the market (Hogan, 1983).

Solutions to this problem are diverse: some say the system of licensing should be eliminated (e.g. Friedman (1962), Baron (1983)), while others say the boards should be reformed with representatives from a lay public or other related professions, or bring the board under strict state supervision (Cohen, 1980).

Dewees continues that licensing a physician to exercise hospital privileges is another method of maintaining competence. According to the authors, hospitals are well situated to screen the medical staff they employ or to whom they grant privileges; through requiring a general license and certification in a specialty with a minimum period of training in university-affiliated programmes, this may be a solution to the above-mentioned problems\(^{12}\). Furthermore, responsibility for the input regulation can be directly imposed on the hospital. A third way of trying to reduce medical malpractice is through institutional competence, both in the form of institutional self-regulation (e.g. through voluntary accreditation) and direct government regulation (e.g. through their establishment and governance etc.)\(^{12}\).

In sum, so far the empirical evidence of adverse or beneficial effects of medical malpractice liability on the performance of services by health care providers is weak. There is some evidence, however, which supports the rent-seeking hypothesis, arguing that physicians would act as an interest group trying to limit the entry into the profession.

### III. Compensation

So far we have addressed the various legal instruments (notably regulation and liability rules) that could be used to prevent medical malpractice. We will turn

\(^{12}\) See D. Dewees/D. Duff/M. Trebilcock (supra fn. 119), 125–127.

\(^{12}\) See D. Dewees/D. Duff/M. Trebilcock (supra fn. 119), 127–128.

\(^{12}\) See D. Dewees/D. Duff/M. Trebilcock (supra fn. 119), 128.
now to the question how, if harm nevertheless occurred, compensation to victims could be awarded. Of course, we will pay particular attention to the question how these different compensation mechanisms in turn affect the prevention of medical malpractice. Hence, the issues of compensation and prevention are interrelated.

When the question of compensation needs to be discussed, there are several possible instruments able to compensate medical malpractice. This chapter will examine the characteristics of the different instruments, and their respective pros and cons will be discussed. The different instruments will be discussed separately, examining if they are often closely connected (through regulation). First we will look at self-insurance (A), then we will look at the working of liability insurance (B). We will also examine interesting alternatives such as first party insurance (C) and compensation funds (D). Finally we will address the question whether victims of medical malpractice should be compensated via tort law or via social security (E).

A. Liability and Self-Insurance

One could first of all trust the system of liability without taking additional measures to guarantee compensation. As has been addressed before in the economic analysis, the compensation function of liability is not so much emphasised as is the function of prevention. Second, this function of compensation of liability rules can only be expected to work in a system of strict liability. Indeed, under the negligence rule, the potential injurer will in principle take due care, and hence not be found liable. Under negligence the victim therefore in principle is not compensated. The victim is also dependent of criteria that lie beyond his reach to get compensation, being for instance, the blameworthiness of the injurer. Hence liability alone does not provide a sufficient guarantee of compensation, given the insolvency risk.

Expanding liability for medical malpractice leads to tensions in the liability insurance market, as we will see later in this chapter. As a result of reduced possibilities to obtain liability cover, some major hospitals have, e.g. in the Netherlands, moved to self-insurance. What often happens is that major (especially public) hospitals self-insure for an important amount and only purchase “excess” insurance for when liability would exceed a specific ceiling. Therefore, in practice a combination of self-insurance and liability insurance, whereby the self-insurance can take the form of a deductible, can be found. Although this combined use of self-insurance and liability insurance may, especially concerning the larger risks, certainly prove to be efficient, one has to warn that a total reliance on self-insurance (meaning that health care providers would not purchase liability insurance any longer) has certain dangers as well.

1. An Alternative for Liability Insurance?

One can really question whether self-insurance should at all be discussed within the scope of a chapter dealing with alternatives to liability insurance,
since the question really arises whether it can honestly be considered as a realistic alternative. The reason we nevertheless wish to mention self-insurance is that it is a form of financial security and that it seems to play an increasing role in medical malpractice, given the uninsurability of certain risks. The reason we are rather sceptical concerning this concept of "self insurance" is that it in fact constitutes a nice word, covering the situation where potentially responsible parties make reserves themselves for potential losses. However, these reserves made by the potential injurer himself cannot be considered "insurance" in the traditional sense, for the simple reason that there is no risk spreading, risk distribution and consequently no loss spreading after an accident happens. Indeed, with self-insurance the risk will not be transferred to insurers, which is typically the feature of most insurance schemes. To be blunt: self-insurance is not an insurance scheme, but a system whereby potentially responsible parties make reserves for future losses.

2. Tax Deductible Reserves

One could therefore question why there is often discussion concerning self-insurance at all if it constitutes nothing more than a private reserve by the potentially responsible party for future losses. The reason is usually a fiscal one. If these reserves were made without any specific goal, they could be considered as hidden profit by tax authorities and could thus be taxed. If, on the other hand, the tax system allows these reserves (and could even encourage them by making them deductible), this self-insurance becomes a way in which the potentially responsible parties could make reserves for future losses in a tax friendly way. Sometimes these reserves are referred to as "captives". These captives, to which there is often reference in the literature, are in some cases again nothing more than reserves made by hospitals to cover future losses.

3. Reserves as Secured Transactions

A positive aspect of these tax devices allowing reserves for future losses to be made is obviously that there may be at least a minimum guarantee that these amounts will be used to cover future medical malpractice. However, from a policy perspective, the administrative authorities controlling e.g. whether minimum financial security may be at hand, will usually be rather cautious with accepting self-insurance as proof of sufficient financial guarantee. The fact that these reserves are made today does indeed not necessarily mean that the amount will still be available at the moment that the loss occurs. This is obviously the case if it is only the potentially responsible party himself (e.g. a hospital) who has the right to decide what will be the ultimate goal of these reserves made in the form of self-insurance. Moreover, even if the reserve were still available if e.g. medical malpractice happens (and the potentially responsible party would hence not have decided to take the money with him to the Bahamas), the question will arise whether the reserve made can effectively be used to cover the damage caused by medical malpractice. In the absence of specific statutory provisions protecting the reserve as a specific security, the reserve will simply be considered as one of the assets of the hospital. If one
4. Advantages and Disadvantages of Self-Insurance

Self-insurance, as described above, obviously has the advantage that it is probably a lot less costly than systems of risk distribution, especially if risks are shifted to an insurance undertaking. The amount paid (via a premium or a contribution) by the potentially responsible party to an insurer or a fund might often be much higher than the actuarially fair value of the risk. The reason is that an insurer usually also charges a sum of money to cover the administrative functioning of its service (so-called loading). Therefore the premium paid is often a lot higher than the objective value of the risk. Thus one can understand why potentially responsible parties might want to self-insure and only purchase excess insurance for high amounts.

An obvious point is that victims still should have the guarantee that a health care provider who is found liable will also have the possibility to be able to pay the compensation due to the victim. Self-insurance is not necessarily a waterproof guarantee against insolvency. This would only be the case if regulation could guarantee that the money set aside for covering medical malpractice is only to be used for that specific goal. Second, liability insurance has the major advantage that a risk spreading via so-called "economies of scale" is possible. A liability insurer has the possibility to bring together similar, but unrelated, risks and can thus increase the expected utility of all insured by reducing their risk aversion. This major benefit of insurance (risk spreading) is obviously lost with self-insurance. Moreover, liability insurance may have the advantage that a specialised insurer (or broker) can acquire accurate information on risk and can thus, via the insurance policy provisions, require specific preventive measures from the health care provider. Efficient insurance policies can thus lead to a reduction of the medical malpractice risk. Finally, a self-insurance of health care providers (especially hospitals) can lead to redistribution problems. Assume that a public hospital would not purchase liability insurance. In that case they would simply run the risk of having to pay major amounts as a result of liabilities and would then pass on the costs to the taxpayer and not necessarily to those who benefit from the services of the health care system. Passing on the risk to the taxpayer is obviously only possible for public hospitals. This could hence also lead to a distortion of competition between private and public hospitals. As a reaction to a failing insurance market, potential injurers have arranged for risk redistributive agreements to share (major) risks together. This development can be seen for example in the Netherlands, where hospitals established together "Medi-Risk".

In sum, self-insurance may be a low cost and useful instrument which may allow the setting aside of assets (making reserves) to cover future losses. However, in order to avoid the risk of externalisation of harm (as a result of insolvency), self-insurance can only be considered an effective financial security if guarantees can be provided through regulation that the reserves set aside will effectively be used for the potential losses for which they are meant. Moreover, the basic problem remains that potentially responsible parties may cause losses that can largely outweigh even the assets which they might have set aside in reserve (under so-called self-insurance). Hence, additional mechanisms will still have to be examined to provide coverage.

B. Liability Insurance

Individuals who engage in a risk bearing activity and who are risk averse will need insurance coverage. The way in which insurance works is well known: the insured will pay a premium to the insurance company, who will in principle agree to cover that risk. He charges a premium, which is based on the probability that an accident with damage of a certain magnitude might occur. Traditionally, providers of health care services have covered their exposure to liability via liability insurance. First we will address the theoretical conditions which have to be fulfilled to guarantee insurability of a certain risk, later on applied to the insurability of the medical practice.

1. Risk Aversion and Insurance

Economists have used the concept of risk aversion to explain that many persons will be averse towards risks with a relatively low probability of occurring, but with a possible large magnitude when they occur. The utilitarian approach with respect to insurance has demonstrated that risk creates a disutility for people with risk aversion. Their utility can be increased in case of loss spreading or if the small probability of a large loss is taken away from the injurer in exchange for the certainty of a small loss. The latter is of course exactly the phenomenon of insurance. The risk averse injurer has a demand for insurance; he prefers the certainty of a small loss (the payment of the insurance premium) whereby the probability of a larger loss is shifted to the insurance company, thereby increasing the utility of the injurer. It is remarkable that, in this utilitarian approach of insurance liability, insurance is in the first place regarded as a means to increase the utility of the risk averse injurer, not so much as a means to protect victims as is sometimes argued by lawyers. The reason an insurance company can take over the risk of the injurer is well known: because of the large number of participants the risk can be spread over a larger group of people. The insurer only has to pay attention that he builds relatively small risk groups in which the premium is as much as possible aligned to the risk of the members of that group. In addition to this utility

---


126 See S. Shavell (supra fn. 12), 190.
based theory of insurance which sees insurance as an instrument to increase the expected utility of risk averse persons through a system of risk spreading, Skogh has argued that insurance may also be used as a device to reduce transactions costs.

2. Uncertainty and "Insurer Ambiguity"

Hence, one can argue that also with respect to medical liability these general principles will apply: insurance can provide protection for risk averse injurers. It should be noted that from an economic perspective liability insurance provides protection for the risk averse injurer and only indirectly for potential victims. By accepting a certain loss in the form of the payment of a premium, the future risk in case an injury would happen can be removed from this injurer. This arrangement will be utility maximising for the particular injurer and may reduce transaction costs since \( ex \ ante \) it has been decided who will intervene in case an injury occurs. The insurance company will be able to provide coverage if he can aggregate similar risks in risk groups and thus spread the risk among his insured via the law of large numbers.

Obviously for every insurance scheme, medical liability insurance included, it is crucial that the insurer possesses accurate information on the likelihood that the event will occur (the probability) and on the possible magnitude of the damage once an accident occurs. This is necessary to make accurate premium calculation possible and to set aside a reserve in case the accident for which insurance coverage was sought occurs. The expectations on probability and magnitude of the loss are essential for the insurer to be able to calculate his so-called actuarially fair premium. With the so-called loading costs (for, among others, administrative expenses) and, depending on the market structure, a profit margin, this will constitute the premium to be paid by the insured. In this respect medical liability insurance is obviously not different from any other type of insurance. The multiplier of these \( p \times D = \text{probability} \times \text{Damage} \) constitutes the actuarially fair premium. The reason the insurer can take over this risk is the law of large numbers: a larger group of insured with a similar risk can be brought together in a risk group and thus risk spreading becomes possible.

If the insurer ideally has \( ex \ ante \) perfect information on the predictability of the probability and the magnitude of the damage, we call the particular risk insurable. It is precisely on the basis of statistics that the insurer will require information on the likelihood that the risk will occur with a particular insured; statistics may also provide information on the possible magnitude of the damage. Both these requirements may, however, be a problem in the case of medical (liability) insurance, while predictability of the liability risk is obviously a crucial element to guarantee the insurability of medical liability. The question

---

therefore arises whether the predictability of the liability risk can be increased, even in the absence of reliable statistics or whether in that case the particular risk should be judged as insurable. The literature has indicated that uncertainty concerning the probability of the damage is of course an element with which the insurer can — in principal — take account *ex ante*. If there is uncertainty, because of a lack of reliable statistics, this should not necessarily lead to the conclusion that a particular risk is uninsurable. We are then dealing with the concept referred to as “insurer ambiguity” addressed by Kunreuther, Hogarth and Meszaros. They argue that the insurer can react to this uncertainty concerning either the probability of the event or the magnitude of the damage by charging a so-called risk premium to account for this unpredictability. Hence, an insurer can in principle also deal with a “hard to predict” event by charging an additional premium. Although theoretically the additional risk premium is hence the answer to insurer ambiguity, in practice the insurer will at least need some information to make more than an educated guess concerning the risk premium he has to charge. Moreover, the fact that an insurer finds himself in a competitive medical liability market may well drive him to take out liability insurance even when an appropriate risk premium cannot be charged. If this is not the case, the insurer might restrain from coverage and uninsurability could arise.

3. Curing “Moral Hazard”: Remedies

A further condition for insurability is that moral hazard can be cured. Moral hazard is the well-known phenomenon that the behaviour of the insured injurer (and every insured for that matter) will change as soon as the risk is removed from him, relating to the fact that the incentives for prevention will be diluted (or at least reduced) as a result of the simple fact that a certain risk is insured. This is precisely the essential contradiction in liability insurance. If risk is fully removed from the injurer and shifted to the insurer, the injurer will indeed not have the incentive for taking care that was given to him by the deterrent effect of having to pay compensation in case of an accident. Marc Pauly has, by the way, indicated that in fact this behaviour of the injurer is not immoral but completely rational since he simply reacts to varying costs for his behaviour.

a) Monitoring

A first best solution is a detailed control of the insured. In that case the premium conditions would be adapted exactly to the behaviour of the insured and the premium would reflect the care taken by the insured. In an optimal world this should give the insured incentives to behave exactly as if no insurance

---

were available and the premium would reflect the true accident risk. Of course this first best solution is only possible in the ideal world where control by the insurance company would be free of charge and information on the behaviour of the insured readily available. In practice this is of course not true. There are, however, some means for a control of the insured and a differentiation of premium conditions is possible according to certain groups of risk. This can either be an ex ante screening with a higher premium for certain high risk groups or an ex post premium increase or change of policy conditions based on previous loss experience. This is the so-called experience rating. Much of insurance legislation is also aimed at reducing moral hazard. Think in this respect e.g. of the prohibition, contained in many insurance laws, to insure accidents which are caused with intent.132

In sum: In order to remedy moral hazard a liability insurer should, as much as (financially) feasible adapt the premium to the individual risk of his insured: rewarding good behaviour (efficient care) and punishing (via premium increases) "bad" behaviour.133 Such a "risk management" is obviously only possible when the insurer possesses accurate information on the risk and thus has the possibility for an effective risk reduction.

b) Exposing the Insured to Risk

A second best solution is exposing the insured partially to risk. This is considered second best because insurance should ideally exactly aim at removing risk from the injurer. Exposing the insured to risk will mean that some degree of risk aversion will remain. This has, on the other hand, the advantage that the insured injurer will still have some incentives for care-taking although he is insured. This exposure to risk can be either at the lower level of damage or at the higher level. One could indeed think of a system with a deductible whereby a lower threshold applies or one could introduce an upper limit on coverage whereby the insured bears his own loss in case the damage exceeds the insured amount.

When the insurer is capable of adequately controlling moral hazard, efficient insurance policies should be able to control the behaviour of the insured and provide incentives for prevention in the same way as tort law does.134 In that case these incentives are no longer provided via tort law but through efficient provisions in the insurance policy. Such a control of moral hazard is obviously possible in the liability insurance for health care provision as well.135

133 See for the basic principles of the control of moral hazard S. Shavell, On moral hazard and insurance, (1979) QJE, 541–562.
135 See S. Shavell (supra fn. 49), 40–42.
c) Combination

In practice one will of course see a combination of both systems of the control of moral hazard. Usually there is some degree of differentiation within the policy conditions, a deductible and an upper limit on coverage. Of course the methods used depend upon the information costs, but also on the value of the insurance policy. Obviously an insurer will more readily tend to invest resources in making a nicely tailored insurance policy for a large company that pays a substantial premium than in case of consumer risks.

If moral hazard is controlled optimally through the use of the above-mentioned devices, the insured will again behave as if no insurance coverage were available, with the benefit that the disutility of risk is removed from him. The incentives for care taking are in that case no longer given by liability law since the threat of having to pay compensation to a victim is shifted to the insurance company. In case of insurance the care taking of the injurer is achieved through an appropriate adaptation of the policy conditions to the behaviour of the individual insured. This also explains why liability insurance has a very important social function. Under liability insurance the insurer has to guarantee that the insured will take efficient care and thus have an incentive to avoid accidents. This makes clear that an appropriate control of moral hazard is not only in the interest of the individual insurer, but also of society. If there were no efficient control of moral hazard, insurance would on the whole do more harm than good. Priest pointed out that the insurance crisis in the United States was, among other reasons, caused because the insurer did not have enough information regarding the insurability of several subjects. Hence, the insurer could not differentiate the risk group enough to control the moral hazard sufficiently.

4. Adverse Selection

Already above we indicated that the insurance is based on a system of loss spreading. Therefore the insurer needs a minimum number of similar risks he insures. At the same time risk pools have to be constructed as narrowly as possible, meaning that the average premium in the risk pool should correspond with the risk of most of the members in the particular pool. Were this not the case, then the average premium would be relatively high for low risk members who would then leave the group. In that case the well-known phenomenon of adverse selection could emerge, which has been described in the seminal paper of Akerlof on the market for lemons.

Adverse selection will, in other words, arise if potentially responsible parties fail to disclose their true risk profile appropriately, which may endanger the

---

137 See G. Priest, [1987] YLI, 1521–1590. We will come back to this American insurance crisis in the next section.
narrowing of risk pools. Adverse selection can therefore only be remedied if the insurer is able to construct the risk pools as narrowly as possible.

5. Risk Differentiation

a) Introduction

The literature has indicated that the appropriate remedy for both moral hazard and adverse selection is risk differentiation. I will first identify what the theoretical benefits of risk differentiation are; then I will show, by using a Dutch example, that there is, in liability insurance in general, but also with respect to medical liability, far more scope for risk differentiation than is used today. Then we will turn to medical liability insurance by showing how, through appropriate specialisation and monitoring, the insurer should be able to check the medical reliability of a caretaker, which constitutes the effective risk differentiation in the medical case.

b) Risk Differentiation: Theory

It follows from the economic principles of liability insurance that an adaptation of the policy conditions to the individual risk is essential to control both moral hazard and adverse selection.

Can such a system of efficient liability insurance be endangered as a result of an increasing liability, which we can notice today in various legal systems? Obviously expanding medical malpractice should not endanger the insurability or the provision of specific health care services. There are obviously remedies, available to liability insurers, to cope with expanding liability. One remedy is an appropriate differentiation of risk. If the insurance policy requires preventive action from the insured party and provides for a corresponding reward in the premium, this should give optimal incentives to the insured for accident reduction. Thus risk pools should be constructed as narrowly as possible so that the premium reflects the risk of the average member of the particular pool. An adequate risk differentiation obviously assumes that an insurer has adequate information on the risk. One of the answers, for instance, in the area of medical malpractice, which will enable a risk differentiation, is a specialisation on medical risks. It is highly important for the liability insurer to be able to reward good risks for preventive action in order to prevent them from leaving the risk pools. In that respect a specialised broker can also play an important role in passing information on risks from the health care provider to the liability insurers.

139 Discuss above in the section on ‘Curing moral hazard’, supra no. 123–127.
140 See for the control of moral hazard within medical insurance, S. Shavell (supra fn. 49), 35–64 (40–42).
141 Above, we, however, noticed that this is apparently the case in some countries. See in that respect the section on ‘defensive medicine’, supra no. 99.
c) Towards an American Insurance Crisis?

George Priest has argued that the insurance crisis in the United States (which plays a major role, also with respect to medical malpractice) has, to a large extent, been caused by the fact that the tort law system was increasingly used as an instrument of victim compensation. He pointed out the fact that the liability insurance crisis in the US was caused by the adverse selection process that results from an information asymmetry between insurer and insured\textsuperscript{143}. Indeed, the insurer often has poor information on the precise qualities of the insured (there is therefore asymmetric information). If the qualities of these insured (whether they are good or bad risks) cannot be signalled to the insurer, the average premium which the insurer will charge will be relatively too high for the "good risks"\textsuperscript{144}. This may lead the good risks to leave the pool and finally to a total unravelling of risk pools. Priest argues that this process has been caused by an increasing use of liability as a compensation mechanism and led to a crisis whereby, for certain services, liability insurance was not available any longer. Moreover, Priest argued that this crisis victimised the lower income groups especially, while premium increases were caused by the higher income groups in particular\textsuperscript{145}.

A further differentiation of the risk is obviously only efficient as long as the marginal benefits of this further differentiation outweigh the marginal costs of such a differentiation\textsuperscript{146}. Risk differentiation certainly does not mean that insurers would have to use an individual tariff in each case\textsuperscript{147}. The possibilities for individual differentiation will inevitably also depend upon the value of the particular insurance policy. For mass insurance products with a low premium, risk differentiation can only take place in general categories.

It seems important to learn from this American example that there are certain dangers in an expansion of medical malpractice from an insurance perspective. This is the more true if, as is the case in many European legal systems, the liability insurer increases premiums as a result of expanding liability, but the health care provider has no possibility to pass these increases to the patients, given the price regulation of services.

\textsuperscript{143} See G. Priest, [1987] YLJ, 1521–1590. Priest has been criticised by Viscusi, who claims that there were reasons for the product liability crisis in the U.S. other than adverse selection on its own (W.K. Viscusi, The Dimensions of the Product Liability Crisis, [1991] JLS, 147–177).

\textsuperscript{144} See on this process of adverse selection G. Akerlof, [1970] QJE, 488–500.

\textsuperscript{145} This is precisely the argument made above that a shift to strict liability will cause a negative redistribution.

\textsuperscript{146} Generally the question is whether the benefits of particularisation outweigh their costs, which has to be addressed in tort law when a standard of care is defined (R. Posner, supra fn. 26)), but also when legal rules are made (I. Ehrlich/ R. Posner, An Economic Analysis of Legal Rule-Making, [1974] JLS, 257) or standards are set (A.I. Ogus, Quantitative Rules and Judicial Decision-Making, in: P. Burrows/C. Veljanovski (eds.), The Economic Approach to Law (1981), 210–225.

d) Significance of Competition for the Insurance Market

The role of brokers can be a very important one, because they can stimulate the competition between insurers by passing on information to clients. The liability insurance will only lead to an efficient solution when the insurance markets are competitive and thus premiums and policy conditions will be nicely tailored to the individual needs and the behaviour of the insured in order to control moral hazard optimally. In practice, however, many restrictions on insurance markets exist\(^{148}\). Important differences remain in that respect between the various European Union Member States. Insurance markets seem to be fairly competitive for instance in the United Kingdom and the Netherlands, but far more concentrated in for instance Germany and Belgium. In other research, the negative consequences of a high concentration on insurance markets with respect to premiums, but also for the incentives of the insurer to control the moral hazard problem have been addressed\(^{149}\). Indeed, if monopolistic premiums can be set, an insurer will have less incentives to align his premiums to the individual behaviour of the insured and thus there is less control of the moral hazard problem. When this is the case, the insurance will lead to an increased risk of damage\(^{150}\). It can hence be concluded that the statement that liability insurance will lead to an efficient solution is a result of the assumption that there is perfect competition within the insurance market.

e) Risk Differentiation: The Medical Malpractice Case

Considering the current practice of many European insurers, especially as far as the liability risk of hospitals is concerned, one is struck by the fact that so little use is made of the possibilities of risk differentiation. Up to now, one lump sum premium has been charged for a whole variety of risks where the premium is sometimes based e.g. on the number of hospital beds. Such a global tariff for a whole variety of different risks obviously does not correspond to the economic need for individual risk differentiation.

This system might have worked in a legal system where liability law was not used as the main source for compensating victims. But now that governments in Western Europe are increasingly withdrawing from social security systems, victims may need to use the tort system more often\(^{151}\). This will inevitably force insurers into more effective risk differentiation. Moreover, the insurer will have to invest in procuring information on the specific risk so as to be able to differentiate. This may trigger a need for specialisation. If an insurer can recognise good risks through his specialised information, he may offer these risks a reduced premium and thus receive a comparative advantage. In-


\(^{151}\) See M. Faure/T. Hartlief (supra fn. 34).
In sum, this shows that insurers may yet learn from the economic argument in favour of risk differentiation. In a world of growing liability this seems to be one of the remedies that allows insurers to compete on the basis of specialized information on risk while at the same time providing an adequate answer to increasing liability. It should be noted, however, that interesting legal questions arise with respect to the limits of risk classification; for example, whether very detailed risk differentiation collides with the non-discrimination principle. Wils argued that too detailed differentiation according to age, gender or sexual preference of the insured could, under certain circumstances, be considered a violation of the non-discrimination principle, laid down in European law. This may not, however, directly cause problems in the field of medical liability where risk differentiation is much needed.

Finally, in the literature concerning liability insurance of medical malpractice, it has been pointed out that the insurer in many legal systems hardly enforces control of moral hazard, because the insurer is of the opinion that he cannot improve the quality of medical treatment. When this would be the case, then the only possibility the insurer would have is to end the insurance-agreement when damage payment would be too large. If the insurer, for whatever reason, decides to break the agreement, there will be no risk differentiation for medical malpractice, and consequently there will be a loss of efficiency.

C. First Party Patient Insurance

1. Introduction

We previously indicated that self-insurance may not constitute the optimal alternative for a well-functioning system of liability insurance for the compensation of damage caused by medical malpractice. However, we just indicated

---

138 This is explained, through the fact that the medical profession is seen as a respectable profession, for which it is not necessary to have external control. See also R. Dingwall/P. Fenin, A respectable Profession. Sociological and Economic Perspectives on the Regulation of Professional Services, [1987] IRLE, 51.
that liability insurance also requires stringent conditions (such as an effective risk differentiation) to function in an optimal way. It may not always be possible to fulfill these conditions since increasing pressures are put on liability insurance, given an expanding medical malpractice insurance in Europe. It is therefore likely that in some countries conditions may be such that traditional insurers no longer provide coverage for liability insurance. This seems to be the case e.g. in the Netherlands. Therefore the question arises whether alternatives are available to compensate damage caused by individual malpractice. Before discussing these alternatives, we should remember that in Europe today a large part of the medical malpractice bill is still primarily paid by the social security system. To a large part liability law therefore either deals with the right of redress of those social security institutions or with damages for pain and suffering of victims. This, therefore, somewhat reduces the need to look for alternatives since victims today are already compensated for a large part of their losses via the social security system. We will come back to the role of social security below in E.

2. Differences Between Liability and Patient Insurance

There is another alternative which may compensate victims of medical malpractice, which is often advocated, being a first-party patient insurance. This phenomenon is extensively debated in the literature. The major difference between liability insurance and first-party insurance is well-known. Liability insurance is a third-party insurance, whereby the insurer covers the risk that his insured (the health care provider) will have to compensate a third party. A “patient” insurance is considered a so-called first-party insurance, since in that case compensation is awarded directly by the insurer to the victim. Whether such a first-party patient insurance can be considered as an efficient alternative for third-party liability cannot be answered in general terms. This depends to a large extent on the details of such a proposal and more particularly on the question whether the patient’s insurance is combined or not with the liability of the health care provider. The underlying principle in a first-party insurance is that the insurance – in principle – pays as soon as damage occurs, provided that the victim can prove that his damage has been caused by the insured risk, irrespective of the fact whether there is liability of a third party. The advantage of a first-party insurance is obviously that the transaction costs are relatively low and that risk differentiation will be a lot easier. The reason is simply that the insurer covers the risk of the patient directly. It is therefore much easier for the patient to signal particular circumstances to the insurer than with a liability insurance. The problem with liability insurance is always that the insurer is insuring the risk that his insured (the potential injurer) will harm a victim (a third party) of which the properties are unknown ex ante to the insurer, whereas under first-party insurance the insurer directly covers the victim. In the ideal world of first-party insurance the insurer directly covers the victim,

155 This argument is especially advanced by G. Priest, The current insurance crisis and modern tort law, [1987] YLJ, 1521–1590.
e.g. the risk. He can therefore monitor the risk directly and in principle provide a much better risk differentiation.

3. Causal Relationship with Medical Treatment Still Required?

A question that arises in the context of first-party patient insurance is whether a causal relationship with a medical treatment is required. If such a causal link would still have to be proven by the patient, a first-party patient insurance might not necessarily be easier than the current third-party liability system. The problem is indeed that the compensation system should not be constructed in such a way that the patient is compensated as soon as he is dissatisfied with the result of the medical treatment\textsuperscript{156}. Therefore, also in a so-called patient insurance scheme it will be necessary to examine whether the damage suffered by the victim is the result of medical treatment\textsuperscript{157}.

4. Financing a First-Party Insurance Scheme

Moreover, also in a first-party insurance scheme the question will have to be answered who finances the system. In principle, it is the patient who finances a first-party insurance. That is obviously a major difference with liability insurance. It will obviously probably be difficult to "sell" a first-party insurance scheme when this has to be financed by the victims. This explains why most patient insurance schemes which either exist (such as e.g. in Sweden) or are discussed in the literature are not first-party insurance in the true sense, but more social security schemes. In those cases, it is the government who finances the insurance scheme. Indeed, since potential victims already pay for their protection against future losses, their willingness to contribute additionally to a first party scheme will probably be low. Obviously one should avoid that a supplementary compensation scheme would only lead to more redress from social security institutions. Combined with liability insurance this may lead to an inefficient cumulation of compensation schemes.

5. Relationship with Liability Insurance

Finally, also under a first-party patient insurance scheme, the relationship to liability law will have to be cleared. One question is whether such patient insurance scheme would replace the liability system. That would therefore amount to a new regime whereby the liability system would be retained. From the patient's perspective this would mean that the patient would receive a compensation with relative certainty (he still needs to prove causation with a medical treatment – which may not be easy – but does not need to prove negligence of the health care provider), although the compensation will usually be

\textsuperscript{156} See for more details H. Koziol (supra fn. 51), 31.

\textsuperscript{157} The administrative costs of such a patient insurance accident scheme are, given this causation requirement, therefore not necessarily lower than the costs of the liability system, so R. Bowles/P.H. Jones, Professional Liability: an economic analysis (1989), 71 and R. Epstein (supra fn. 86), 257–258. This is, however, denied by T. Vanswevel, De Civielrechtelijke Aansprakelijkheid van de Geneesheer en het Ziekenhuis (3rd edn. 1997), 875–876.
lower than the full compensation awarded in tort law. From the health care providers' perspective, the question arises how they would still have incentives for prevention without a liability system. Many have warned that a true insurance scheme without liability for medical malpractice would dilute the incentives for care of the health care providers.\footnote{158} If, on the other hand, the liability system continues to exist in addition to the patient insurance, this would lead to a cumulation of insurance, which can probably hardly be considered as efficient.\footnote{159}

Moreover, the question arises what amounts would be paid under first-party coverage? Usually the amounts paid in first-party insurance are lower than the full compensation, which is in principle awarded under tort law.

\section*{D. Compensation Fund}

\subsection*{1. Introduction}

A related discussion, concerning no-fault compensation schemes is based, not on the idea of a first-party patient insurance, but on the idea of the installment of a compensation fund for victims of medical malpractice. The tendency to install compensation funds instead of using liability law has also reached the area of medical malpractice.\footnote{160} The reasons advanced for such a compensation fund vary, but it is often argued that the traditional liability system with liability insurance is not able to provide full victim compensation and is therefore defective in that respect.\footnote{161}

\subsection*{2. Compensation Funds: Necessary?}

Moreover, compensation funds are sometimes advanced with the argument that patients would otherwise receive no compensation. This argument often neglects the fact that, especially in Western Europe, social security systems already cover e.g. most of the medical expenses and the income loss. The un-

\footnote{158} This has been argued by H. Koziol (supra fn. 51), 32 and by P. Danzon, [1990] GPRI, 3–22 (13).

\footnote{159} So H. Koziol (supra fn. 51), 31.

\footnote{160} The idea was also launched in France to set up a ‘Fonds d’indemnisation des victimes d’accidents médicaux graves survenus en l’absence de faute des soignants’, [17 February 2000] Le Monde, 19.

compensated part is therefore usually the pain and suffering and the top of the income\textsuperscript{162}. One can therefore question whether far-reaching solutions, such as the instalment of a compensation fund, are necessary to deal with this relative marginal problem of the uncompensated damage. Providing a compensation fund for specific victims of accidents and not for others necessarily implies a selection which, once more, leads to inequality\textsuperscript{163}.

Because of these – and other – arguments one can understand that many scholars are relatively critical concerning proposals to introduce compensation funds to cover medical malpractice risks. Those proposals almost always encounter severe criticism from economists\textsuperscript{164}.

I shall now discuss the different aspects of medical compensation funds.

3. Various Funds

One cannot escape the impression that often – especially at the political level – funds are advocated as a miracle solution for all problems of medical malpractice whereby no clear definition is given of the specific funds. I will briefly sketch the kind of funds that might play a role when medical malpractice is discussed. This short overview will make clear that usually very different goals are achieved by these various funds that play a role in accident law today.

a) Limitation Fund

In American literature the idea of a fund is sometimes used to refer to the situation whereby mass damage is caused by similar products or services. The problem that usually arises in these cases of serial damage (e.g. with certain medicine – DES) is that the liable enterprise may be willing to agree to a settlement with the victims on the condition that he can offer a certain sum to all the victims whereby he can reach a final settlement for the damage caused by the specific tort. In a certain sense the manufacturer then raises a fund that will have to be used to compensate all victims. It could be called a limitation fund since the enterprise agreeing to such a settlement usually wishes to limit its liability to the amount brought into the fund. In this case no risk spreading with other (potential) manufacturers takes place, since only the liable manufacturer finances the fund. As reward for the payment of a particular sum the duty to compensate is usually limited to the amount made available: a limitation fund. The goal of a limitation fund in these cases is to have an adequate instrument to divide the available proceeds among the victims in case of serial damage\textsuperscript{165}.

\textsuperscript{162} See H. Koziol (supra fn. 51), 30.
\textsuperscript{163} So H. Koziol (supra fn. 51), 33.
\textsuperscript{164} See e.g. the critical comments concerning fund proposals of R. Epstein (supra fn. 86), 257–267.
\textsuperscript{165} This was the reason that the European Directive on Product Liability also incorporated the idea of a limitation fund in art. 16 which gives Member States the option to limit liability in case of serial damage to an amount which may not be less than 70 million ECU. However, only few Member States have used this option.
This may play a role for damage caused by e.g. pharmaceuticals, but less in case of medical malpractice. In that case there will rarely be serial damage.

b) Advancement Fund

A second arrangement referred to is the advancement fund. The prepayment or advancement construction is a remedy for long-lasting civil procedures concerning liability and insurance coverage issues that can last much longer than the life of the victim. An example is the case of asbestos victims in which it has been argued that it is highly unfair that (relatives of) victims only receive compensation post mortem because of the relatively short time between the discovery of the illness and their death. Therefore an advancement fund has been advanced in the Netherlands as a remedy for asbestos victims. The reason why compensation funds are sometimes defended with respect to medical malpractice is not so much related to long-lasting procedures, but more to the uninsurability of certain risks or because it may be difficult for the victim to receive compensation via tort law. If these are the problems to be solved, an advancement fund is no adequate remedy.

c) Guarantee Fund

Guarantee funds are well known as instruments to protect victims against the possible insolvency of a liable injurer or his insurer. The advantage of a guarantee fund is that it only intervenes for the so-called excess risk, being the risk for which in the specific case for various reasons no insurance coverage is available. However, it is essential that a guarantee fund only intervenes if other compensation mechanisms, such as insurance, have failed. The possible role of a guarantee fund for compensating medical malpractice will be discussed below.

d) A General Medical Malpractice Fund

Another alternative could be a medical fund that would generally operate as a substitute for liability and insurance. However, if these compensation funds were to be used as an alternative for the liability system, inevitably the question will arise how they can be financed in an adequate way. Therefore I shall now focus on basic differences between compensation funds on the one hand and liability combined with insurance on the other hand as devices to provide compensation for damage. In this respect it will first be argued that the general ideas of economic analysis (e.g. that liability gives incentives for prevention) should also play a role if compensation is regulated via instruments other than tort law.

---

4. General Principles of Fair and Efficient Compensation

No matter how a compensation system is organised, the incentives for prevention of medical malpractice should always remain untouched. They are important when it comes to judging the effectiveness of alternative compensation mechanisms. In this chapter it has often been stressed that liability rules do not just have a compensating, but also a preventive effect. Liability rules can only have a preventive effect if the duty to compensate is put on the one who actually contributed to the risk. This means that a duty to compensate should in principle only rest upon the one who actually contributed to the risk.

A second related principle is that this duty to contribute should also be related to the amount in which the specific treatment or hospital contributed to the risk. This principle is usually automatically respected in liability law via the causation requirement. The duty to compensate under tort law is indeed usually limited to the damage that the specific tortfeasor himself caused. However, also if a collectivisation of the compensation takes place it remains important to guarantee that the tortfeasor only contributes financially in relation to the amount to which he contributed to the risk as well. This is reflected in insurance policies in the idea of risk differentiation. It simply means that bad risks pay a higher premium than good risks. This principle should also be applied if a compensation fund is installed. This remains important since it will give incentives for prevention to the contributors to the fund. Bad risks will be punished and good risks should be rewarded.

Moreover, when fund solutions are discussed, very often the question who should finance such a fund is neglected. In principle only those who actually contributed to the risk should contribute to the fund. However, when it is clear who (e.g. which hospital) has contributed to specific risks, traditional liability law could be applied as well. Moreover, as we will argue below, the major difficulty of applying tort law to medical malpractice, being the causation issue, will also arise if one wishes to set up a fund efficiently.

An alternative would obviously be to have the compensation fund financed by the State. In that case the fund would be financed by the general taxpayers. But in that case one could ask the question at the policy level why specific victims of specific accidents (in this particular case of medical malpractice) would deserve a better treatment than other victims.

These principles are not only important from an efficiency point of view (providing optimal incentives for prevention), but also include a fairness element. Indeed, if these principles were not followed, it would mean that good risks would have to pay for the bad risks as well and would therefore in fact subsidise bad risks. This negative redistribution should be avoided and therefore the

167 This idea has been elaborated in no. 136–139.
168 See in that respect critical remarks of Kozioł, who argues that such a separate regime is contrary to the equality principle (H. Kozioł (supra fn. 51), 31).
compensation mechanism, fund or insurance should be financed principally by the ones who really contributed to the damage.

In sum, the compensation mechanism should aim at a differentiation of the contributions due. This differentiation is only possible if the insurance company or agency administering the fund also possesses information on the amount to which the specific activity contributed to the risk. Moreover, the system should be organised in such a manner that the fund only intervenes for damage caused as a result of medical malpractice. Hence, also in a fund solution the causation requirement remains important. It obviously requires a lot of information to determine whether the damage was actually caused by malpractice (and not e.g. by a natural cause). One key element to determine the choice between insurance or funds is therefore who possesses the best information to control the risk and to establish causation.

5. Funds versus Insurance

a) Risk Differentiation

Applying the principles discussed above there are not many reasons why, if both a compensation fund and liability insurance are – in theory – available, a compensation fund would provide better protection against insolvency of the health care provider than the insurance markets (under the condition that there is perfect competition within the insurance market). One could assume that an insurer is better able to differentiate risks since an insurer is specialised in risk differentiation and risk spreading. Insurers therefore possess techniques to determine in what way their insured contribute to the risk. Obviously this assumes that the insurance markets are competitive. In the absence of competition in insurance markets, either the supply of insurance coverage could be too limited or premiums could be excessively high, which could justify a preference for a compensation fund. In this case preference should be given to governmental administrators to manage a medical compensation fund and to monitor risks adequately. The question arises whether the administrator of a compensation fund will better be able to manage the medical malpractice risk than traditional insurance markets, as it is assumed that insurers are better able to deal with classic insurance problems such as moral hazard and adverse selection. The crucial question is indeed whether the administrator of the fund is also able to recognise good and bad risks and hence, to carry through a risk differentiation in the same way a liability insurer would do. Insurance markets have proven to be able to provide adequate coverage for liability and it is not at all certain that an adequate risk differentiation could be performed more effectively by a (public) compensation fund. One cannot see as a matter of principle why a government agency that would run a compensation fund would have better information on risks than an insurer. This might, however,

158

159

158

159

160 S. Shavell (supra fn. 49), 40.
be different if highly technical risks are involved where care takers or hospitals are in a much better position than the insurance company to monitor each other. This would be an argument in favour of e.g. risk distribution among health care providers who can mutually monitor each other. Thus one can understand why in certain countries mutuals have appeared (such as MediRisk in the Netherlands) which constitute an alternative for liability insurance via a form of mutual risk sharing which is different to the (public) funds, discussed in this section.

160 In sum, if both insurance and compensation funds are available, there are no clear reasons why a fund would be the preferred solution. There may, however, be reasons why insurance may not provide coverage for certain risks. In that case funds cannot be compared with insurance since insurance is no alternative.

b) Costs

161 Comparing insurance with compensation funds one should also address the comparative costs of both instruments. Insurance will generally be cheaper because liability insurance policies are not concluded for one activity, but for a whole set of risks. There is hence one insurance policy with transaction costs that are incurred once and an administrative structure within an insurance company that will be forced to an adequate cost reduction by competitive pressures. The costs of risk spreading might also be lower with an insurance company (or with mutual risk sharing) than with a compensation fund. Insurers are indeed specialised in methods for acquiring information on differentiation of risks. In addition, it has been argued in the literature that insurance provides for a reduction of transaction costs between contracting parties, because parties can agree ex ante on a distribution of risks and losses in case of an incident 171. The comparison will obviously also depend upon the type of compensation fund under discussion. In most cases one immediately thinks of a compensation fund run by a regulatory authority. If that is the case, one can of course refer to the literature on the negative effects of bureaucracies to argue that such a publicly operated compensation fund should not necessarily provide compensation at lower costs than the private insurance market. This can be reduced if the fund is administered privately, but in that case a competition with other funds has to be organised to provide incentives for cost reduction. One could only think of cost advantages of a fund if one did not refer to a fund as one that would replace the liability and insurance system, but would be a limitation fund or a guarantee fund, such as the ones discussed above.

6. Examples

162 Although one should, for the above-mentioned reasons, not expect miracles from no-fault compensation schemes for medical malpractice, and there may even be adverse effects as far as incentives are concerned, one can neverthe-

171 This argument has been made by G. Skogb, [1989] JRI, 726–732.
less point at various proposals to introduce these no-fault compensation schemes. Proposals in this direction are discussed in the literature of many of the legal systems. It is probably interesting to point at a couple of examples. This will merely be an introductory sketch. These existing compensating funds for medical malpractice will of course be discussed in further detail in the comparative research. Let us briefly address a few examples of compensation funds to cover damage caused by medical malpractice.

a) Belgian Academic Draft of a Medical Compensation System

The first one is a Belgian academic draft, made by professors who argue that such a fund is necessary because numerous victims of medical malpractice receive no full compensation today. Besides that, care-takers themselves are very unsatisfied with recent developments in the administration of law, together with the fact that liability insurance alone would lead to serious losses for insurers. This last remark of course does not prove that insurance losses as a result of this administration of law are inefficient. Moreover one can argue that the Belgian insurers probably did not bear in mind the legal implications of expanding medical malpractice, and as a result of this made inaccurate calculations. The Belgian professors have therefore proposed a compensation fund, which should be financed by the insurers and by the State.\footnote{See on this fund among others J.L. Fagnart, La réparation des accidents médicaux, in: T. Van-sweevelt (ed.), Responsabilité et Accidents Médicaux (1996), 53–92. This volume contains several contributions in which also the various interest groups (medical profession and insurers) react on the proposal for a compensation fund.} It is, however, unclear, whether the fund will exist in combination with liability law and how a risk differentiation can be applied as far as the financing of the fund is concerned.

b) Compensation Fund of Vienna, Austria

This applies to some extent as well for another proposal, which has not been elaborated in great detail yet, but is nevertheless interesting. It concerns a proposal made by the city of Vienna to erect a compensation fund for damage suffered by patients.\footnote{Announcements on this fund could be found among others in [8 October 1997] Der Standard and in the [8 October 1997] Kurier Wien.} According to this proposal, patients who become victim of medical malpractice in a hospital of the city of Vienna would be compensated through a fund (to a limited amount!), whereby the city of Vienna would be subrogated in the rights of the patient against the tortfeasor. But this does not restrain the victim from submitting a claim via his liability insurance.

This proposal, however, raises various questions. First of all it can be questioned if such an offer can really serve its purpose, that is to avoid “American situations” because the victim still has the possibility to submit a claim for full compensation to a court. So it is clear that it can never replace liability law (it is hardly feasible that the city of Vienna could change Austrian tort law). Second, it remains unclear how the fund should be financed. Since it would be...
impossible to limit compensation to citizens of Vienna because of reasons of equality and the prohibition of discrimination, a financing through the taxpayers of Vienna would amount to a redistribution in favour of non-citizens.

c) The Swedish Model

Many argue that the Swedish patient insurance system for medical malpractice is a success story, which should be copied by many other legal systems. Earlier studies have shown that it is doubtful whether the Swedish system does indeed provide the success story that foreigners often want to see in it. First of all, one may not forget that Sweden already had a far-reaching social security scheme. In addition, the Swedish Act on Patient's Damages was only introduced in 1997 and is based on a compulsory insurance, to be taken out by health care providers to the benefit of patients. The system provides limited amounts of compensation. Patients can claim damages on a no-fault basis, but still have to prove a causal connection with the medical measures. Hence, this Swedish model looks more like a first-party insurance than a fund solution. Moreover, in the Swedish system, victims can still claim damages in tort, based on the Swedish Tort Act of 1972. The fact that liability law remains in existence, in addition to the no-fault compensation scheme may therefore still provide sufficient incentives for prevention. Moreover, the system is still so recent that there is no conclusive empirical evidence yet as to either the financial viability of the system or its effect on care taken by the health care providers. It seems therefore too soon to draw any final conclusions yet from this Swedish experience.

7. A Compensation Fund for Medical Malpractice

So far I have sketched that in general there are very few reasons to expect that a compensation fund would provide better compensation than a private insurance market. Problems of risk differentiation and the need to prove causal link will arise as well in fund solutions. Nevertheless a fund solution may be interesting if liability insurance were unavailable, e.g. because of the unwillingness of insurers to provide coverage. Which kind of compensation fund would provide the best solution and under which conditions?

A first option is to implement a compensation fund to guarantee compensation in case of insolveny of the injurer or his insurer. In that case a fund does not replace the liability and insurance system, but only intervenes in a particular case when the injurer or his insurer was found insolvent. This combined use of the liability system, insurance and a guarantee fund for the insolvency risk has

174 See e.g. the referral to the Swedish system in Austria by J. Pichler Rechtsentwicklungen zu einer verschuldnahunabhängigen Entschädigung im Medizinbereich (1994) and J. Pichler, (supra fn. 161), 35–46 and in Belgium by T. Vansweevelt (supra fn. 157), 866–872.
175 See M. Faure (supra fn. 44), 304–305.
176 See the discussion in the previous section.
the advantage that the incentives of the liability system will remain untouched
and that the fund will only have to intervene in the event of insolvency.

However, the question arises whether such a fund could at all be used sepa-
rately from compulsory insurance (in the broadest sense). Compulsory insur-
ance may be necessary to protect the innocent victim, and give the duty to a
potential injurer to purchase liability insurance to protect the victim against in-
solvency of the injurer. Indeed, if insolvency of the injurer is the problem, it
seems more logical to discuss the introduction of a duty to insure instead of
immediately advancing a fund-solution. If one looks at the use of guarantee
funds now, one can note that they are usually used to provide for the cases that
an insurance company goes bankrupt. However, it is important to stress that if
one wishes to provide a protection against the insolvency of the injurer, it is
more appropriate to focus on a variety of financial mechanisms, discussed ear-
erlier in this section, that force a potential injurer *ex ante* to provide for compen-
sation for future losses instead of focusing on an *ex post* fund solution. Gener-
ally, one can doubt whether it is useful to focus on a guarantee fund that
should e.g. intervene in case of insolvency of an insurer instead of focusing on
other *ex ante* compensation mechanisms to provide coverage.

Another option would be to abolish indeed liability and (traditional) insurance
altogether for damage caused through medical malpractice and to replace it
with a fund. First of all one would then have to argue that there are good rea-
sons to believe that the liability and insurance system will not be able to pro-
vide adequate compensation to victims of medical malpractice. In addition at-
tention will then have to be paid as well to the question whether a fund is able
to solve the problems found in a liability and insurance system. Moreover, the
fund will then have to be structured in such a manner that only those who ac-
tually cause the risk have to contribute to the fund (optimal risk differentiation
for optimal prevention) and that only those victims whose losses are caused by
medical malpractice are compensated by the fund.

Obviously these issues cannot be answered generally or merely at the theoret-
cal level. In this introductory chapter we merely wanted to address a few is-
ssues which, from an economic perspective, will play a role in judging the ef-
fectiveness of a general compensation fund for medical malpractice. It is
therefore important to examine how these issues have been addressed in the
various legal systems where such a collective compensation scheme has been
proposed or introduced. This will be the subject of the comparative research.
E. Social Security or Tort Law¹⁷⁷

1. Introduction

A final option to compensate medical malpractice is through a system of social security. The system of social security can be seen as a system where limited compensation as a way where minimum compensation can be reached and also, depending on the legal system, the legal foundation, e.g. illness or disablement, is mostly of no importance. In practice in many states social security plays a major role concerning compensation of victims of medical malpractice. For example additional medical expenses or loss of income can be compensated by the social security, irrespective of the reason for the illness. It is important for a social security system, contrary to a liability insurance system, that the preconditions to enter a social security system will be relatively low, but also that the amounts to compensate are limited. Social security will compensate for a minimum amount, and not, as is the case with a liability insurance system, a full compensation of the victims. Social security is therefore more "Existenzsicherung."¹⁸¹

2. Difference Between Tort Law and Social Security

The basis of tort law, and in that respect it is clearly different to social security, is that the occurrence of damage as such is not enough. Even recently the Dutch Supreme Court clearly stated that:

"The mere fact that a certain act took place which resulted in damage for another party does not necessarily imply that a claim on compensation exists on the basis of tort."¹⁸²

Some other element needs to be added to grant the victim a right for compensation of his damage. Traditionally this was "fault" although there is a clear tendency in many legal systems towards strict liability.¹⁸³ In the literature it is claimed that usually some shortcoming on the side of the injurer is required as basis for a liability in tort.¹⁸⁴ A basic rule in tort law is that if the victim can

¹⁷⁷ See more general on the relationship between social security and tort law from an economic perspective M. Faure/T. Hartlief, Social security versus tort law as instruments to compensate personal injuries: a Dutch and economics perspective, in: U. Magnus (ed.), The impact of social security law on tort law (2003), 222-255; hereafter these general notions are applied to the specific situation of victims of medical malpractice.


¹⁷⁹ See H. Koziol (supra fn. 51), 30.


¹⁸¹ See H. Koziol (supra fn. 51), 22.


¹⁸³ The tendency from fault to strict liability has also been addressed on economic grounds in the previous paper on strict liability (M. Faure (supra fn. 3)).

prove that his loss was caused by this wrongful (in the sense of shortcoming) act of the inurer, he can in principle claim full compensation, both of pecuniary and of non-pecuniary loss (although the amount and scope may obviously differ in the legal systems), on the basis of tort law. This full compensation would only be reduced (to some extent) if the victim himself also contributed to the loss. The essential ideas of liability law are therefore that the victim has a claim on full compensation of his loss if a causal relationship can be proven with a wrongful act.\(^{185}\)

The starting point of social security is different: social security provides, as has been stated in the previous section, compensation (e.g. of lost income or health care expenses) irrespective of the cause of the illness or disability. The social security system traditionally only covers pecuniary loss and even then usually limited in time and as far as the amount is concerned. Contributory negligence of the claimant in principle does not exclude a claim on compensation in social security law. It needs to be mentioned that contributory negligence almost never plays a role in social security systems, whereas in most tort law systems contributory negligence of the victim may lead to a reduction of the claim on the part of the victim.

Fault traditionally does not play a role in social security systems: a victim can claim compensation under social security without having the need to prove the fault of a third party. That is different in tort law, where traditionally a fault, but at least a wrongful act will have to be proven. Moreover, under social security the victim will only need to prove a certain condition (e.g. illness or unemployment) and that this condition meets the statutory requirements for compensation. It is not necessary to prove a causal link between this condition, which gave rise to a loss and the act of a third party, which is typically a requirement of tort law. Moreover, in social security the financial situation of the victim might play a role in deciding upon the amount of compensation. These redistributitional elements are, again, contrary to the basic approach of tort law: a victim is, if the specific conditions of tort are met, entitled to full compensation, irrespective of his personal wealth.\(^{186}\)

A starting point of social security is also the idea of solidarity between the citizens. All citizens contribute (in accordance with their income) to the social security system. This contrasts rather sharply with tort law, which does not know any principle of solidarity between inurers or victims. Only within (liability) insurance is it sometimes said that there is solidarity between all the insured. That, however, is a different type of solidarity than the one in social security. The solidarity referred to in insurance is only a system of risk spreading whereby all insured (via premium payments) contribute to the loss when the risk occurs with one particular insured. However, since premium is in principle risk based in insurance, bad risks should pay more than good risks. Risk

\(^{185}\) T. Hartlief (supra fn. 180), 26–28.

\(^{186}\) T. Hartlief (supra fn. 180), 28–29.
differentiation is therefore the leading principle in insurance, not the type of income dependence solidarity, which is known in social security.\footnote{187}

178 Trying to identify a dividing line between tort law and social security law, one could very generally state that from an economic perspective the general function of tort law is prevention of accidents, whereas the general goal of social security law is providing compensation for the damage caused as a result of injuries. A major difference following from this dividing line is that tort law is triggered as a result of an accident whereby a third party is held liable for having caused the injury, whereas this is not at all necessary under social security. Social security is therefore a system whereby compensation is, if the legal conditions are met, provided directly to the protected victim, irrespective of whether his injury was caused by a liable third party. It may well be that social security intervenes in cases where it was entirely the victim’s fault that injury was caused\footnote{188}. However, also in social security prevention of harm is increasingly stressed as being important as well\footnote{189}. However, the payments made under social security schemes cannot be considered as having mainly a prevention goal. Traditional social security still focuses on compensation of losses\footnote{190}.

3. Combined Use in Practice

179 Although the starting points are therefore different (torts: deterrence; social security: compensation) the economic literature has indicated that if one had pure systems relying totally on either tort or social security, the systems would have to take care of compensation issues as well as prevention of accidents. We mean the following: even in a system which would primarily rely on torts, one will notice that the tort system (which economically serves the goal of deterrence) will be complemented with systems of insurance to meet the second goal (compensation). Thus a potential tortfeasor could take out liability insurance coverage which might (if the injurer is found liable) serve the interests of the victims as well, since it provides a guarantee against the insolvency of the potential injurer. Moreover, under a pure tort system victims (who would fear that they would receive no compensation) could choose to take out first-party insurance coverage to cover for personal injury losses. This is, by the way, what obviously happens in practice a lot: many potential victims will take out

\footnote{187} This is further developed in M. Faure, The applicability of the principles of private insurance to social healthcare insurance seen from a law and economics perspective, [1998] GPRI, 265–293.

\footnote{188} This principle that social security provides compensation irrespective of the behaviour of the beneficiary may well be different in some legal systems; one can imagine cases where compensation under social security is denied if e.g. the injury was caused intentionally.


\footnote{190} S. Klosse (supra fn. 189), 8–9.
accident insurance coverage to provide (additional) coverage in case personal injury happens to them. Hence, the deterrence oriented tort system might be combined with (liability or first party) insurance schemes to serve the goal of compensation\textsuperscript{191}.

The same applies, however, as well for a system whereby automatic compensation (irrespective of the behaviour of the beneficiary) would take place under social security. If such a perfectly working social security compensation mechanism would exist whereby all potential victims would be guaranteed compensation for their personal injuries, some other systems than tort (assuming that this would not exist) would be needed to guarantee that those who may have an influence on the accident risk behave properly. The systems, which are then advocated to be combined with social security, are systems of safety regulation. Thus, the model then assumes that victims are compensated via social security, whereby this is combined with \textit{ex ante} government regulation, which is enforced via administrative or criminal law. Regulation, not tort law, then serves the goal of deterrence\textsuperscript{192}.

This is a model which has been advocated by many economists: they have often argued that if society wishes to reach the goals of both deterrence of accidents and compensation of victims for personal injury, it could choose between on the one hand a tort system (for deterrence) + insurance systems (for compensation) or on the other hand a system of regulation (for deterrence) + social security (for compensation). This combination of private and public regulation of safety has especially been advocated by Skogh\textsuperscript{193}. In practice of course a combined variety of all of these instruments of deterrence and compensation exists, although the focus on each of the particular instruments may change per society and over time. However, all of the systems (torts, insurance, regulation and social security) are interrelated and mutually influence each other. Hence, one can expect that e.g. if the government would withdraw from the compensation of personal injury via social security, victims would be forced to make an increasing use of tort law and insurance to meet this same goal.

4. An Example of Social Security

In the Netherlands in case of medical malpractice traditionally social security provided a limited compensation of lost income and took care of expenses for the health care system. This corresponded with the traditional view that it is social security, which takes care of the "Existenzsicherung"\textsuperscript{194}. In addition to

\textsuperscript{191} See T. Hartleif (supra fn. 180), 28.
\textsuperscript{194} See H. Koziol (supra fn. 51), 21–35.
the social security system in the Netherlands tort law traditionally played a role, but only a modest one. Victims only used tort law to receive compensation for the top of their income (the part which was not covered by social security) and to get coverage for non-pecuniary loss. These were precisely the types of damage not covered under social security. Prevention was to a large extent guaranteed through regulation, imposing specific safety duties on the health care provider.

183 This resulted in a system whereby victims primarily received compensation via the social security system to provide some “Existenzsicherung”. Tort law could be used, if the specific conditions were met, for that part of the damage, which was not covered under social security. Obviously some interrelationships between those systems could exist in the sense that (again, under specific conditions) the social security system might use tort law to attempt to get recovery of the amounts paid to the victim.

184 It is striking that in this traditional (Dutch) (but to some extent European) system tort law was a luxury system and at the same time of rather limited importance in the compensation of victims. Indeed, only a limited amount of the damage which occurs in society is covered via tort law.\textsuperscript{195} The largest part of damage was covered either via social security or via private first party insurance. Tort law can be considered a luxury system in the sense that it provides a guarantee of in principle full compensation for the damage suffered and even compensation for non-pecuniary loss. That is a luxury, so it has been held in the literature, which the social security system cannot afford\textsuperscript{196}. Indeed, the essence of an “Existenzsicherung” is that it provides a minimum, but not the “luxury” of full compensation. The economic reason why the social security system cannot guarantee full compensation (including compensation for non-pecuniary losses) are manifold, the costs of full recovery would be high and would lead to higher premiums or an increased pressure on public budgets. In addition non-pecuniary loss will be different for every individual, whereas social security usually works with more or less fixed, at least standardised compensations.

5. Tort Law: Luxury or “Existenzsicherung”?

185 Therefore it was held that this “luxury” of tort law can only be provided in exceptional circumstances and when specific conditions are met. Hence, tort law cannot guarantee full compensation to every victim of personal injury, not even to every victim of an accident. It will in most systems be dependent on the behaviour of the injurer whether the victim can obtain this full compensation. That again corresponds with the economic insights sketched out above that tort law is, at least in systems where it does not have to provide this social

\textsuperscript{195} This has also been proven empirically (by calculating the specific contribution of tort law compared to social security) by A.R. Bloemenbergen, De invloed van verzekeringen, in: Schade lijden en schade dragen (1980), 16–17.

\textsuperscript{196} T. Hartlief (supra fn. 180), 29–30.
ance a debate is going on today whether the principles of risk differentiation can be carried through in all respects. Cousy has often advocated that insurance companies should take ethical principles into account as well. These may lead them to accept some solidarity between good and bad risks. These ethical principles in other words impose certain limits on a very detailed risk differentiation. Wils has also pointed out that a detailed risk differentiation may conflict with basic human rights if, for example, risks are differentiated according to gender or sexual preferences. A basic difference between social and private insurance is obviously that the policy maker wants to achieve income redistributive goals through the financing of the health care system. This, however, obviously may have a price as far as the efficiency of the system is concerned.

Today many also advocate a combination of limited social insurance schemes, providing basic medical treatment, financed by income dependant premiums with the possibility of purchasing additional insurance on the market according to the preferences and demands of the particular individual. Hence, it is not always necessary to make a strict choice between private or social (health care) insurance, but a combination of systems is possible. In that respect one could obviously think of a compulsory social insurance scheme for the larger risks, where the above discussed model hazard problem does not play an important role, and the possibility of private insurance for the smaller risks.

7. No-Fault Compensation Scheme

The basic difference between the function of tort law and social security is also relevant for the discussion on a no-fault compensation scheme for medical malpractice. Indeed, although the idea of a general no-fault compensation scheme may seem like an attractive solution to guarantee some compensation to victims and protect health care providers against liability claims, many doubt that these solutions can be realised in practice. It fits into a general tendency to desire to compensate all accident victims. This tendency of legal doctrine and policy makers to seek full compensation for victims, also of medical malpractice, is obviously not an isolated phenomenon. One can notice in various areas of accident law that victims increasingly refuse to accept less than full compensation. The old adagium “the loss lies where it falls” apparently is


no longer accepted. This has led to various tendencies in accident law and also in the field of medical malpractice. One tendency is to expand the scope of liability by lowering the standard of liability.\footnote{This obviously has to do with the desire of each victim to shift his loss to someone else (see H. Koziol (supra fn. 51), 21).}

A second tendency is that the traditional difference between tort law and social security seems to become smaller. The traditional idea was that social security would provide for an easy compensation (with a low threshold), but also for a limited amount, whereas full compensation could only be awarded when the more complex conditions of liability under tort law were met.\footnote{See H. Koziol (supra fn. 51), 33–34.} Victims of accidents now seek “the best of both worlds”: they seek the low threshold for compensation of social security, to be combined with full compensation under tort law. This combination will, however, inevitably lead to problems at the insurance level. Traditional tort law and liability insurance have not been developed as mechanisms which should guarantee full compensation to all victims of accidents.\footnote{See T. Hartleff (supra fn. 180), 56–57.} One problem is that tort law is now expanded or combined with pleas in favour of no-fault compensation schemes whereby often the question who should finance this expanded protection seems to be neglected.

The future will probably show that victims and policymakers will have to make a choice between either an automatic compensation, which can be warranted through no-fault compensation schemes (social security, first-party insurance or compensation fund), but then the damages awarded are necessarily limited (and the question still has to be answered how incentives for prevention can be given) or to rely still on tort law with its full compensation for a necessarily limited number of victims.

\section*{F. Empirical Evidence\footnote{This section is adopted from the analysis of D. Dewees/D. Duff/M. Trebilcock (supra fn. 119), 136–146.}}

Let us, to conclude this discussion of the compensation aspect of medical malpractice, also briefly address the available empirical literature on the efficiency of no-fault compensation systems, as this has been summarized by Dewees, Duff and Trebilcock.

It is clear from the no-fault compensation systems existing in New Zealand, Sweden and Finland, that in order to be eligible for compensation through such system, there are certain conditions a patient claiming to have become disabled or injured through medical malpractice needs to fulfil. The systems are usually confronted with three “causation criteria” in order to determine if compensation is justified, being\footnote{D. Dewees/D. Duff/M. Trebilcock (supra fn. 119), 137–138.}
1. intention and expectation
2. failure to intervene
3. probabilistic and proportional causation

195 Ad 1.

Many medical treatments can have adverse effects on the patient, being the result of that particular medical treatment (e.g. loss of hair with chemotherapy). However, it is not sufficient that the patient’s disability be caused by medical intervention, it must also be the unintended or unexpected result of such treatment (P. Weller, 1993). These cases must be excluded from the compensation scheme, to avoid becoming a general insurance disability scheme however the injury was caused209.

196 Ad 2.

It is not necessary that the injury is caused by active medical intervention; it may also occur when a disability is caused or aggravated by a failure to prevent or minimise the patient’s condition at a stage when it was medically possible to do so, with the exception that the condition was medically impossible to diagnose before the condition was merely untreatable206.

197 Ad 3.

The iatrogenic injury is often typically one element in a complex causal chain leading to the patient’s disability, where the disability can be attributed to both medical causation and the patient’s underlying condition. So the definition and proof of a compensable event involve difficult issues of probabilistic and proportional causation. Therefore, many proposals for comprehensive no-fault patient compensation plans specifically limit eligibility to cases in which medical care is both a probable cause of the adverse outcome (Ross and Rosenthal et al. (1975)) and a significant or material cause of resulting disability207.

According to Dewees the empirical evidence of the no-fault compensation schemes in New Zealand, Sweden and Finland is very limited208. They discuss the effects of the no-fault compensation schemes in New Zealand and Sweden by mentioning the frequency of compensation, benefits paid, the disposition of claims, and the experience of insurance costs under each plan209. Some of this information is interesting, although it is only partial and e.g. does not discuss the effects of a no-fault compensation scheme on the incentives for prevention.

New Zealand: Concerning the frequency in which patients are compensated after medical malpractice, figures indicate that considerably more injured

209 The statistics and conclusions for this empirical evidence of this section are all adopted from the analysis done by D. Dewees (supra fn. 119), 144–146.
patients are compensated by the accident compensation scheme than obtained benefits through the tort system existing before the system of compensation in New Zealand. But, it is not said that all patients are compensated, mainly because a patient needs to be able to prove that the treatment is the predominant cause and therefore fails in the claim for compensation. About only 60% of "medical misadventure" claims are ultimately granted by the New Zealand Accident Compensation Corporation (Report of the British Medical Association, 1983).

Dewees continues that since 1992 a reform of the compensation scheme of medical malpractice has taken place. No compensation was awarded anymore for non-pecuniary losses, as well as establishment of a deductible to exclude minor or short-term losses, to decrease unnecessary payments of compensation.

Dewees mentions about disposition that reports prior to the 1992 reforms indicate that most claims for "medical misadventure" were granted with little delay (British Medical Association, 1983) and that administrative costs accounted for less than 10% of total premium dollars collected under the New Zealand scheme (Hodge, 1983). It needs to be mentioned that in some cases this was quite the opposite in that delays and administrative costs were substantial.

_Sweden_: The Swedish Patient Insurance Plan also had an increase in the number of compensation claims for injuries. But here as well, one notices that through restrictive eligible criteria for compensation, a percentage of between 50 and 60% of all claims filed are eligible for compensation. Dewees mentions as well that most claims are resolved quickly and with little expense (Oldertz, 1984), 80% within 7 months (Brahams, 1988), in comparison with the United States on average 3 years, with 45% more than 2 years (US GAO claims, 1987). And insurance costs have been quite economical in Sweden under this plan, compared to the U.S., and are still nowhere near the amount paid in the U.S.

### IV. Directions for Further Research

#### A. Importance of Law and Economics

In this first part some economic observations were provided concerning optimal prevention and compensation of medical malpractice. We tried to provide an introduction to this study on the possibility of introducing a no-fault accident scheme for damage caused by medical malpractice in the Netherlands by pointing at the insights which follow from the economic literature. We believe it is highly useful to start this research project with an economic analysis, particularly since the law and economics methodology has addressed the goals of liability and insurance. Indeed, if attention is given, as in the current project, to possible alternatives for liability and insurance (such as a no-fault compensation scheme) it seems important that first some ideas are formulated concerning the functions of liability and insurance.
Obviously we do not claim at the normative level that law and economics is the only methodology which could provide an answer to the question what the functions of liability and insurance should be. The advantage of law and economics is, however, that in this literature the effectiveness of various legal instruments to compensate damage caused by medical malpractice has been extensively studied. For that particular reason we started this research project with an overview of the law and economics literature, also since we wanted to avoid “reinventing the wheel”. It seems useful to take into account the state of the art of the law and economics literature for the remainder of this study.

These economics findings do not provide final answers and will have to be supplemented inter alia with the conclusions from the comparative research. However, this law and economics part could probably have some relevance for the comparative analysis as well. Therefore we tried in this first part to identify some of the questions which are considered relevant by economists when it comes to designing an optimal system for the compensation of damage caused by medical malpractice. It will obviously be most interesting to compare these first results of the law and economics analysis with the outcomes of the comparative analysis.

Let us now focus on some of the first results of the economic analysis and indicate how they can be implemented in the remainder of this research project.

B. Prevention via Various Instruments

First of all it is important to stress that law and economics apparently make a clear distinction between prevention and compensation. Economists stress that prevention is the best way of victim compensation and that is no doubt also true for victims of medical malpractice. Therefore economists also point at the fact that various (no-fault) compensation mechanisms may also have an effect on the prevention of medical malpractice. These potential preventive effects should of course be taken into account when the efficacy of a no-fault compensation scheme is finally judged.

The economic literature which was summarised in the second chapter made clear that economists usually distinguish two sets of legal rules which could reach the goal of prevention of medical malpractice. One possibility is to rely simply on traditional tort law; the other one is to use regulation to prevent medical malpractice.

1. Liability for Medical Malpractice

For economists liability rules have an important deterrent effect: the foresight of being held liable ex post will induce the health care provider ex ante to take efficient care. This is, according to economists, the case both for a negligence rule or for a strict liability rule. A strict liability rule has, however, the advantage that it shifts the decision upon the optimal way to reduce risk to the health care provider. Indeed, a negligence rule works perfectly (in the sense of pro-
viding correct incentives for prevention) only if the judge can adequately set a standard of due care. In cases (such as medical care) where highly technical risks are involved and one can assume that the potential injurer (the health care provider) has better information on how to prevent risks than the judge, a strict liability rule might provide better incentives. However, strict liability is efficient only if the potential injurer (the health care provider) has sufficient funds at stake to compensate for the damage he might cause. In case of an insolvency risk it is therefore essential that some mechanism is introduced to cure that particular risk. As we will discuss below, insurance might be such a remedy.

Although we showed that, from a theoretical perspective, strict liability might be an appropriate rule for medical malpractice (provided that the insolvency risk can be cured) we should be careful with generalising this result. A discussion of the Coase theorem showed that in cases were the potential injurer is able to pass on the costs of an increased liability to consumers or patients, a shift towards strict liability will have no distributive advantages. Moreover, there might even be negative re-distributional effects in case of strict liability since particularly the higher income groups will benefit from this increased protection. Whether this passing of increased liabilities is always possible in the medical malpractice context is not easy to answer. In some cases an increased liability might lead to higher prices for medical services; that would mean that patients (or their health care insurers) would pay for an increased protection that is awarded to them. In other cases regulation might prohibit an increase of the tariffs for health care, which would mean that the health care provider himself remains stuck with the effects of an increased liability. Nevertheless, the lessons of the Coase theorem remain important in the context of medical malpractice. Some economists (the papers were discussed above) have even advocated that the whole medical malpractice crisis could be solved through contractual arrangements whereby patients would ex ante bargain on the required level of care and the eventual compensation in case of medical malpractice. Such a differentiated arrangement of care and compensation is obviously the type of division of risks which are advocated by the Coase theorem.

2. Regulation of the Level of Care

Although economists stress the importance of liability to deter medical malpractice, tort law is certainly not the only legal instrument which provides incentives to the health care provider for an appropriate prevention of medical malpractice. In practice preventive measures in the medical sphere result largely from regulation. We explained, using Shavell’s criteria for safety regulation, that there are indeed important arguments in favour of a regulation of medical care. In many cases it can be assumed that the judge (or the parties involved) are not able to assess the optimal health care adequately. Moreover, more particularly in cases of medical malpractice, there is a huge risk that a patient would never discover that his damage has been the result of medical
malpractice. Indeed, a major problem to make liability rules work in case of medical malpractice is obviously the requirement of causation. That is of course the major weakness in any system of liability or compensation for medical malpractice. Since patients who demand medical services usually already have a weak health condition before the medical services are provided, it will often be very difficult to prove that a deterioration of the health condition has been the result of medical malpractice. Since there is, hence, a major probability that, even if there were a case of medical malpractice, health care providers will escape a liability suit, it can certainly be argued that tort law may not have a sufficient deterrent effect to prevent medical malpractice. That explains the existence of regulation.

However, regulation of medical care in practice is not (as is the case in other areas of risk, such as the environment) the subject of extensive public regulation. In many countries the standard of medical care is set by professional organisations, which have received these powers from public authorities. For the remainder of this study it remains obviously important to also address the question whether these standards of care and the control mechanisms within medical law can be considered as adequate tools of quality control. This question is particularly important since the economic literature discussed in this part indicated that although the appropriate medical care standard should be set in regulation, liability rules could still have an important additional deterrent effect. It was therefore held that merely complying with regulations should not free a health care provider from liability.

An important question within the context of the possible introduction of a no-fault compensation scheme (and a possible abolition of liability) is therefore whether the (self-) regulatory and other mechanisms to control quality in medical law can be considered so adequate that in practice the supplementary deterrent effect of liability law is no longer needed. That is obviously an argument which is often advanced by the defenders of no-fault compensation mechanisms. The argument goes that liability law should not have a deterrent effect; the quality of medical care should be guaranteed via mechanisms in medical law and compensation should be awarded via a no-fault compensation scheme. Whether liability law can indeed still have an additional deterrent effect is, however, an issue, which cannot be answered in the abstract and which will be addressed in the comparative analysis.

C. Compensation

1. Various Possibilities

The third chapter in this law and economics analysis addressed a variety of possible compensation mechanisms. It should be mentioned that these were only presented as possible mechanisms which can provide compensation for damage caused by medical malpractice and which might have some importance for the remainder of this project. However, other arrangements which
were not discussed yet in chapter 3 are of course possible. This is for instance the case with risk distribution arrangements between health care providers. These risk distribution arrangements, which are neither traditional liability insurance nor no-fault compensation funds, already exist today in the Netherlands and are therefore certainly of importance.

The discussion of compensation mechanisms made clear that today a variety of compensation mechanisms exists and that in practice various mechanisms also intervene in the compensation of damage caused by medical malpractice. It is obviously an important issue for the remainder of this study whether all of these mechanisms should be replaced by a no-fault compensation fund or whether such a fund (which is precisely the subject of this research project) should only have a supplementary function.

2. Self-Insurance

Indeed, we noticed that today compensation for damage caused by medical malpractice is to some extent based on a mechanism which is politely referred to as "self-insurance", but which in fact is nothing more than a health care provider (physician or hospital) who covers a part of the risk himself without taking liability insurance. We indicated that although the risk of this arrangement is obviously that is does not solve the insolvency problem, there are advantages as well. One clear advantage is that the potential injurer can cover a large part of the potential losses with his own assets and can hence only purchase liability insurance according to his own demand for coverage. Thus self-insurance could lower the total insurance costs. Hence, when a no-fault insurance scheme is finally discussed at the policy level the extent to which health care providers will still be allowed to self-insure should not be neglected.

3. Liability Insurance

The same is true for the other traditional compensation mechanism, being liability insurance. In the discussion concerning compensation we identified some elements which may affect the insurability of medical malpractice. It is well-known that especially in the U.S. insurers have claimed that medical malpractice became an uninsurable risk as a result of the liability crisis which arose in the late eighties. Apparently also in other markets traditional insurers have increasingly withdrawn from the medical liability insurance market. For the remainder of this research project it is important to identify what, according to insurers, specific problems would exist concerning the insurance of medical liability risks. This is an issue which will explicitly be addressed in the comparative analysis. If medical malpractice were uninsurable via traditional liability insurance, that obviously constitutes an important argument to look for alternatives.
4. First Party Insurance

One of these alternatives is what is sometimes referred to as a patient insurance scheme. In the economic literature several advantages are advanced of a first party (victim) insurance scheme, compared to third party (liability) insurance schemes. However, notwithstanding the major advantage that first party insurance would better enable a risk differentiation (since the insurer directly insures the risk of the particular patient) still a lot of questions arise. One of the questions is under what kind of conditions the first party patients will provide compensation. In that respect it has to be stressed that the need to prove a causal link between the damage of the patient and medical malpractice will still exist. In addition the question arises how such a first party insurance scheme will be financed and finally the question arises whether liability law will still be applied in addition to such a first party insurance scheme.

Although many issues remain therefore unresolved, the possibility of a first party patient insurance is definitely an interesting alternative which merits careful examination, also in the context of a no-fault compensation mechanism. Therefore, the comparative analysis will also examine whether these first party mechanisms exist and their functioning will be critically examined.

5. No-Fault Compensation Scheme

A well-known alternative, which is the central focus of this research project is of course the no-fault compensation scheme for medical malpractice. We have briefly discussed this possibility in the context of the economic literature concerning compensation funds. The economic literature has indicated that there might be a case for such a compensation fund in cases where alternatives (such as liability insurance, first party insurance or risk distribution arrangements) would not work. However, the economic literature has also warned that one should analyse very carefully what the specific reasons are for the uninsurability of medical malpractice. It might indeed well be that the problems, which cause an uninsurability of medical malpractice will also lead to problems if a no-fault compensation scheme is installed.

Indeed, economic analysis points at the fact that if a fund solution is organised, this should in principle be arranged in such a way that only those who contribute to the risk should contribute to the fund and also to the extent to which they contributed to the fund. In other words: since the contribution to the fund should be differentiated according to risk this should have a preventive function. If one installed a general compensation fund which would e.g. be paid by the general taxpayer (without any risk differentiation), the question would of course arise whether such a system would have a negative influence on the incentives for prevention. In addition, a no-fault compensation system should also be organised in such a manner that only those patients whose damage is caused by medical malpractice could benefit from the fund. Hence, for a no-fault compensation fund to work efficiently, a causation requirement will have to be introduced, just as this is the case in liability law.
Again it will be the subject of the comparative analysis to see how the systems which have such a no-fault compensation fund deal with these important issues of risk differentiation and causation.

6. Social Security

Finally the discussion on the differences between social security and tort law pointed at the fact that within the context of a discussion on a no-fault compensation system we cannot neglect the question what type of damage the compensation fund should pay. Indeed, in most western legal systems (such as is the case in the Netherlands) a large part of the damage suffered by a victim of medical malpractice is compensated for by the social security system. In fact victims traditionally only used tort law for the non-compensated part of their income and for pain and suffering. The question will obviously have to be answered whether a possible no-fault compensation fund should become a general fund, which would compensate all damage caused by medical malpractice (the latter choice would also have important financial implications) or whether the no-fault compensation fund should have to compensate those heads of damage which have not been compensated under the social security scheme. At the same time the question will have to be analysed whether a no-fault compensation system for medical malpractice should provide the same “luxury” as tort law, being a full compensation of all damage and a compensation for non-pecuniary losses. Most systems outside of tort law, which provide “automatic” compensation to victims contain important restrictions on the amounts of damages. Hence, the question will also have to be addressed whether the victim who will be compensated through a no-fault compensation scheme will have to pay a price for the automatic compensation e.g. in the form of financial caps on the amounts.

Some of these issues will undoubtedly be addressed in the comparative analysis, more particularly by the countries which are familiar with a no-fault compensation scheme.

D. Further Issues

Finally it should be stressed that this first part only generally focused on what law and economics has to say about prevention and compensation of damage caused by medical malpractice, without having the ambition to answer all questions that could be raised in that respect. The goal was mainly to use the economic analysis as an analytical tool to identify some of the issues which could play a role, at least from an economic perspective, when a no-fault compensation scheme is introduced. However, we of course could not test the validity of the assumptions of the economic analysis yet. That was not the purpose of this first part. We provided some information on empirical studies with respect to the effects of various legal rules on compensation of medical malpractice as a background. But many other devices related to liability and insurance could be examined to guarantee an optimal compensation of medical malpractice. This has, on the one hand, to do with the structure of liability
law. One could for instance discuss the necessity of financial caps or of statutes of limitation. On the other hand one could also examine techniques to increase the insurability of the medical risk, such as the system of coverage in time (loss occurrence or claims made coverage). Finally, a policy maker wishing to regulate medical malpractice in the public interest might also want to consider solutions for the insolvency risk, which have often been advocated in the literature, such as a duty on the health care provider to provide some kind of financial security to remedy the insolvency risk. Since, however, the main focus of this research project is the feasibility of a no-fault compensation scheme these other issues have not been elaborated on at this stage. They might be considered, however, if it would appear e.g. that they play a considerable role in some of the legal systems which will be analysed in the comparative analysis. Indeed, in a further stage of this project the findings from the comparative analysis will be compared with the economic principles sketched in this part.

V. Bibliography


A. Bloembergen, De invloed van verzekeringen, in: Schade lijden en schade dragen (1980), 16-17.


M. Gimpel-Hinterregger, Grundfragen der Umwelthaftung (1994).
E. Mackaay, Economics of Information and the Law (1982).
S. Rose-Ackerman, Rethinking the Progressive Agenda, the Reform of the American Regulatory State (1992).


