

Balancing rationalities: gatekeeping in health care

Dick L Willems *University of Amsterdam, Amsterdam, the Netherlands*

Abstract

Physicians are increasingly confronted with the consequences of allocation policies. In several countries, physicians have been assigned a gatekeeper role for secondary health care. Many ethicists oppose this assignment for several reasons, concentrating on the harm the intrusion of societal arguments would inflict on doctor-patient relations. It is argued that these arguments rest on a distinction of spheres of values and of rationality, without taking into account the mixing of values and rationalities that takes place in everyday medical practice. If medical practice, then, does not follow a single, pure rationality, can it also incorporate the societal rationality of the gatekeeper role? Using a case from general practice, I try to show how physicians may integrate societal arguments into their practice in a morally acceptable way. A version of the model of reflective equilibrium and especially Beauchamp and Childress's safeguards, may be helpful both to analyse and teach such balancing of values and rationalities.

(*Journal of Medical Ethics* 2001;27:25–29)

Keywords: Allocation; gatekeeper; physician-patient relations; justice; balancing; ethics

1. Introduction

Gatekeeping by physicians has been both advocated and contested as a way to make health care choices at the level of the individual. A gatekeeper is “a defined point of entry each time care is needed for a health problem”.¹ The role of the general practitioner as an “entry point” has been discussed recently by, among other bodies, the World Health Organization.² The necessity of gatekeeping is usually underscored with three types of arguments: the need to ensure that patients receive *appropriate* care, the need for *budget restraints*, and the need for *justice* in distributing care. During the last few years, normative issues around the gatekeeper role have increasingly received attention.

The term gatekeeping has military origins.³ Among the—allegedly—gluttonous patients who storm the gates of the health care fortress,^{4,5} the gatekeeper selects those who will enter and those who will not. Thus, the gatekeeper metaphor focuses on restraining people from overusing health care, not on ushering in the underdemanding and underserved; also, the military gatekeeper's duty is over once he/she has let the proper people in—he/she has no business guiding them through the system or making sure they leave it when that would be appropriate.

Thus, the gatekeeper metaphor obscures the important question of how to prevent people from getting forever lost in (high-tech) health care. These origins and limitations notwithstanding, many physicians in various countries have accepted both the role and the designation.^{1,6,7}

Different specialties are functioning as gatekeepers: a recent study from the United States showed that 36% were in general internal medicine, 26% in internal medicine subspecialties, 23% in general paediatrics, 7% in paediatric subspecialties, and 5% in family medicine.⁸ In nine out of 15 European countries, primary care physicians have been given a role as either complete or partial gatekeepers for specialist forms of care.¹ In the Netherlands, general practitioners have been gatekeepers for several decades.

Many physicians and ethicists have expressed concerns about the acceptability of gatekeeping by physicians. It is often argued that the patient should, in the end, be the master, not society^{9,10} and that the duty of physicians is exclusively to individual patients, not populations.^{11,12} In gatekeeping, some say, the trust-based relationships between patient and physician risk being eroded by the physician's economic position and tasks.^{13–21} The criticisms seem to depend on Michael Walzer's idea that there are different “spheres of justice” which need to be kept distinguished.²²

On the other hand, many ethicists have argued, at least conditionally, in favour of physician gatekeeping.^{5,23–29} Haavi Morreim mentions a core argument in favour of involving non-medical considerations in the consultation room: as citizens, doctors and patients have to assume a responsibility for health care costs, and if they don't, others will, the result of which will probably be worse.²⁹ Among many other points, Morreim stresses the importance of discussing societal aspects such as costs with patients—such openness could, in her view, limit the danger of losing the trust of patients.

The ethical discussion about gatekeeper systems has concentrated, up to now, on its acceptability, and much less on the conditions under which the gatekeeper role could be fulfilled, if at all, in an acceptable way. Since gatekeeping in some form is accepted and performed by many physicians, the latter question merits more discussion. How do and how should physicians perform the “balancing act”, as Haavi Morreim has called it?²⁹

In order to attack that question, I will need to expand Walzer's “spheres of justice” into “spheres

of rationality". Using the term "justice" in the context of gatekeeper discussions would unduly restrict the analysis, because, as will be shown, justice is one of the rationalities or logics that physicians need to coordinate and balance when they are gatekeepers. Elements from other "spheres", such as needs and efficiency also play a role in many, if not all patient-physician contacts. Whereas Walzer suggests that the different spheres of justice need to be kept apart to avoid one sphere from dominating the others, we see that in practice judgments from different spheres need to be made commensurable and "weighable". The crucial problem with gatekeeping could be that a shared measure is absent and needs to be forged: where the logic of clinical judgment is radically different from the logics of cost-effectiveness and justice, a bridge must be made.

For these reasons, I will use "different rationalities", as an umbrella term for the different spheres of value, justice and goodness ("moralities" also suggested itself, but risks restricting the argument unduly to ethical aspects); however, I will focus on the coordination (in the dictionary sense of putting in the same order or rank and bringing into a common action, movement, or condition) of the rationalities instead of their separation; in that respect the analysis is close to that of Thévenot and Boltanski, who stress the importance (and fragility), in the everyday resolution of problems of justice, of "compromises" between "worlds", as they call it.³⁰ I will specifically address the methodology of reflective equilibrium as a possible coordination and balancing strategy.

The paper proceeds as follows: I will illustrate the practice of gatekeeping with an everyday case from my own general practice; then, I will translate the methodology of reflective equilibrium from ethical theory-building to actual practice; the paper ends with suggestions about good gatekeeping.

2. Gatekeeping and its rationalities—an example

One morning, Mr Jansen walks into the office with some difficulty, right hand against his back, painful grin. He sits down uneasily and complains that he woke up very early with increasing pain in his lower back, the same problem as a year ago. Because of his responsible job in the insurance business he can't miss a day at his work. He has already been trying painkillers for some days, but to no avail, and he asks me to send him to an orthopaedic surgeon right away.

I take his history and examine his back. This leads to a diagnosis of muscular stress and I ask Mr Jansen why he wanted a referral. "Because I don't want to wait as long as last time—that just takes too long."

"This kind of backache", I answer, "just needs time to heal, whatever you do, so I propose watchful waiting. There is absolutely no advantage in an early referral—all our guidelines, which are based on evidence, say so."

Mr Jansen is not convinced by guidelines and tells me that he would gladly be the exception that guidelines usually provide for. He wouldn't ask for a referral if he didn't need it.

Patient preferences are, or should provide, important guidance in the allocation of health care. Many relatively unproblematic consultations indeed follow a "grocery" scheme: patient asks, doctor delivers.

But not all. Physicians often do not and should not automatically consider *prima facie* preferences to be real, or well-informed preferences. They explore, gauge and test the wish of the patient. After all, Mr Jansen's demand for referral might only have been a way of stressing the seriousness of his backache.

But even if Mr Jansen's preference to be referred had been well informed, the physician might not consider it *appropriate* to act accordingly. Appropriateness means that care should be effective—although definitions of effectiveness may vary considerably³¹, medically necessary, and neither too much nor too little. Inappropriate, that is ineffective, unnecessary, overgenerous or frugal care could well be useless or risky or both, but would be irrational in any "sphere". Crucially, patients, in this framework, are not considered capable of deciding without gatekeeper help what care is necessary and effective.

It is important to realise that acting rationally in any sphere involves not only a state of mind, but demands social, psychological, and material resources and skills. In any sphere, if you want to act rationally you not only need to think rationally according to the rules of that sphere, but you also need the material possibility to do what is demanded by that sphere. Thus, making a distinction between appropriate and inappropriate care demands lots of technical resources: sufficient diagnostic skills and technology, and a body of knowledge or guidelines, to mention but a few.

Mr Jansen and I go on to talk about the question whether one should try an expensive option (seeing a specialist) that has a slim chance of benefit and a greater chance of iatrogenic harm, or prefer a less expensive option, even if slightly less effective.

Mr Jansen: "What counts for me is that I get rid of this backache as quickly as possible, even if that is expensive—staying away from work is expensive, too."

I mention the approximate prices of different strategies: "We are really talking about marginally greater benefits, you know. Referral is a lot more expensive and probably no better. You can spare yourself the time and energy, and the surgeon's as well."

Again, Mr Jansen says that whether costs outweigh benefits is something he would like to decide for himself, and he does not feel convinced they do.

The arguments raised in this part of the consultation come from a different sphere—to use Walzer's

term again—and invoke another rationality, within which one should opt for the cheapest effective care (instead of, for instance, care by the doctor with the warmest manner, or in the hospital which is nearest). Appropriateness, at this moment, gets linked to money and resource use.

This is the rationality of cost-effectiveness: referral is irrational, firstly, in the trivial sense that the same benefit could be bought cheaper. Secondly, and less trivially, a referral is irrational if it is too costly in view of the benefits following from it. Thus, a cost-effectiveness logic typically brings in ideas about what an improvement or forestalled deterioration is worth. What do you do if one treatment is much cheaper and at the same time slightly less effective, for instance, if a painkiller would make Mr Jansen's pain bearable, but not entirely cure it, whereas physiotherapy would make it disappear? Can physician and patient be expected to opt for the cheaper treatment? Clearly, this type of balancing occurs often in everyday consultations, but it is not understood how and what its limits are.

What type of resources does a physician need to be able to act cost-effectively? Firstly, he should have at least some idea of prices of diagnostic and therapeutic procedures. Secondly, he should have alternatives available: watchful waiting instead of immediate treatment. Thirdly, physicians need experience in weighing costs against benefits and helping patients to do so, both formally by using some sort of instrument to show the value of treatment options, and informally in deliberating with the patient.

Mr Jansen remains unconvinced. “The last time I had this, it took two months before it went away, as you will recall; referral may be worth a try at least—if it doesn't help, it won't hurt!”

I then raise the idea that Mr Jansen should also give it a chance to recover on its own because the orthopaedic surgeons in the region have long waiting lists, and giving priority to Mr Jansen would mean disadvantaging others.

This argument seems more convincing to Mr Jansen than the others. At the end, we agree that he will try for a few days with painkillers, but if the pain doesn't disappear in a week, he will come back to re-discuss referral.

Here, a third type of rationality comes in: choosing care that does not stand in the way of the needs of others. The guiding idea here is that care should be assigned to people in a just way.³² Now it may be argued that this is a problem for policy-makers in the first place, but justice cannot be kept out of individual care relationships. For instance, it plays a role any time physicians prioritise between patients. Giving priority to one patient in one's waiting room will keep the others waiting. Thus, the physician-gatekeeper is not only dealing with the individual, but with a population of (more or less) similar patients: to some extent, other patients *always* make their way into the sacrosanct privacy of the consultation room. In Mr Jansen's case, the waiting lists

actually exist, but even if there were none, justice would demand that care should primarily be assigned to those who are most in need of it. In the rationality of just distributions, it is the *difference in urgency* that counts.

Justice as an argument demands from patients that they be prepared to step back from their own problem and not consider it as all-important—they should be prepared to waive or at least qualify their presumed right to care in the interest of others more in need of it. It constructs, in other words, patients as citizens. Being rational in this sense again demands resources, for example, information about the number of similar patients and the length of waiting lists. It also demands that general practitioners have the means to weigh the relative urgency of different patients, which is clearly more difficult. It also demands some sort of agreement between gatekeeper and specialist as to what constitutes good reason for skirting waiting lists.

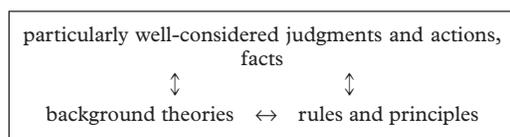
What Mr Jansen's relatively simple case shows is that a physician brings in different types of rationality, or value systems, or types of goodness, with their associated criteria for good and bad practice (jumping from one Walzerian sphere to another). It is, however, unclear how these rationalities are brought into a relationship and made commensurable, how they are weighed, balanced, and coordinated, even if only temporarily. It is especially unclear what indicates that a situation of balance has been reached.

3. Mixing spheres until the equilibrium holds

Putting walls around spheres of rationality could make the work of physicians more difficult rather than help them. What gatekeeping asks physicians and patients to do in many (if not all) consultations, is to coordinate at least the rationalities mentioned above and to obtain a provisional balance between them. The difficult question is: “what is a proper mix of reasons and what isn't?” It can be argued that physicians use some implicit “mixing and balancing method”—a set of rules and skills similar to those indicating how one should mix eggs and mustard to make mayonnaise “hold”.

“Balancing” has been a central concept in ethics since the introduction, by John Rawls, of the reflective equilibrium (RE) model. It has been developed as a methodological instrument for ethical theory development, in order to obtain a coherent ethical theory that is sensitive to the facts of (moral) life.^{33,34} Its main foe is applied ethics which tries to “plug facts into principles”, as Daniels says.³⁵ In the “wide” version Daniels proposes, RE looks somewhat like grounded-theory approaches in the social sciences: from a particular considered judgment in a particular situation, try to develop a more general rule and to link that both to other practical judgments and to a higher level background theory. Then go back to similar situations as the one in which the original judgment occurred—if the rules, principles and theories don't seem to be of any use there, *all* levels may be independently questioned

and adapted until equilibrium (the “theoretical saturation” of grounded theory) is reached. That is the “reflective” part of the model: difficulties at one level reflect on all the others.



The most appealing, but insufficiently exploited, element of the methodology may be its central metaphor: *equilibrium*. “Equilibrium” points to the often unstable state in which a variety of equivalent elements temporarily compensate for each other. Equilibrium is non-hierarchical, it is flat. In other words: equilibrium is a radically *democratic* metaphor. Even the tiniest shift in, for instance, diagnostic results can bring along a disruption of the equilibrium, analogous to Lorenz’s butterfly, the standard example used in chaos theory to show how a tiny local change may disrupt a global meteorological equilibrium (in that sense, a true equilibrium is *always* reflective).

Although RE has been developed for the purpose of developing ethical *theories*, several moral philosophers have argued that the strength of the reflective equilibrium model is that it is very close to how people actually discuss moral issues and could therefore be helpful in discussing *practical cases*. Reaching moral consensus on a particular case, according to them, involves balancing various considered judgments and principles. Even so, this still concerns the level of *debating* a morally problematic case. Is what a physician and a patient *do* in any sense comparable to the balancing act of moral philosophers discussing a case?

I don’t want to suggest that primary care consultations do or should look like ethical debates—heaven forbid!—but still, they have various elements in common with RE: just like ethicists constructing theory, Mr Jansen and I move from considered, but contestable judgments (“all facts considered, Mr Jansen’s back pain should be left alone”) to principles or guidelines (“uncomplicated back pain of short duration does not need specialist intervention”) or directly to background theories (“no treatment for self-limiting diseases”) and back. Consultations involve many heterogeneous elements in which, on all levels, modifications may prove necessary during (a number of) consultations—they “need not exclude anything”, as Van der Burg and Van Willigenburg say about RE.³³ Walzerian spheres are happily, and maybe sometimes naïvely, mixed and balanced. More often than not, the balance reached is as shaky and temporary as the equilibrium metaphor suggests—that is why, at the end of consultations, physicians frequently advise patients to come back a few days later. So maybe physicians have something to gain from RE, and the other way around.

One of the criticisms raised of RE is that it offers no criterion to distinguish good coherence from

bad—reliance on coherence alone may lead to totally abject, but perfectly coherent theories (such as the mob’s set of principles). Another is that RE does not tell us how to coordinate incommensurable rationalities and values to enable us to balance them. Thomas Nagel has told us not to despair and urges an elaboration of Aristotle’s concept of practical wisdom.³⁶ Beauchamp and Childress, in their discussion of RE, seem to take up this challenge by proposing five “safeguards” to guarantee that balancing, ie giving more weight to one norm than to another (“infringement”) is justified:

1. Better reasons support the overriding norm.
2. The moral objective justifying the infringement has a realistic prospect of achievement.
3. No morally preferable alternative actions.
4. The form of infringement selected is the least possible, commensurate with achieving the primary goal of the action.
5. Minimise the negative effects of the infringement.³⁷

One might want to add at least the following:

6. There has been discussion of the infringement between the actors involved (physician and patient).

I don’t want to suggest that such safeguards will solve coordination problems once and for all. Tragic situations of unresolvable conflict will surely remain, but safeguards like these might make it possible to distinguish acceptable forms of balancing rationalities from unacceptable ones in at least a number of cases. Let us see if they help in evaluating the remainder of Mr Jansen’s case.

4. An equilibrium for Mr Jansen

One week later, Mr Jansen does indeed present himself. His backache has been somewhat better for a few days, but when he lifted up his cat he felt it shooting back, and it was exactly the same now as it was a week ago. He reminds me of our previous agreement.

The examination leads to exactly the same diagnosis. I tell him nothing has changed and that medically speaking there is no more reason to refer him than last week. However, I do feel a bit more sympathetic to Mr Jansen’s argument that another week off work will not only cost a lot of money, but also make him feel more miserable. Moreover, Mr Jansen tells me that he has called a few hospitals in the region and has heard that there is one orthopaedic surgeon who has no waiting list at all—so referring him will not harm others.

I feel torn between two possibilities: either stick to the evidence-based medical argument that referral is still unnecessary, or give in because non-medical arguments begin to favour referral. After a while, explaining to him that it is only for his reassurance, and that I see little benefit in it myself, I refer Mr Jansen to the orthopaedic surgeon he has proposed.

What has made me change my mind is that the balance of reasons has shifted: waiting lists have disappeared, making justice rationality irrelevant; budget rationality has begun to veer towards referring as a

cost-effective option, and even appropriateness has begun to point to referral, if only because it might assuage Mr Jansen's fears.

Was this good or bad balancing? Does the set of criteria mentioned above help to decide that question? Quite roughly, the case seems to conform to most of the "safeguards"; at the same time, it shows their restrictions. Other elements might make this consultation an example of how to balance: firstly, there is openness on the part of the physician to changes in weight of the different factors. Furthermore, balancing is a cooperative effort: both the physician and Mr Jansen are involved in it. Both physician and patient display a degree of balancing skill—it's a skill that can be learned. But possibly the most important element is that there was no evidence, at least in this case, of waning trust.

5. Conclusion

Throughout the paper, I have stressed the resources and tools needed to balance values and rationalities. Apart from balancing skills, physicians need tools and resources consisting of specific information (about prices and waiting lists, for instance), but also of diagnostic machinery, of guidelines, and of time, to mention but a few. Where these tools are lacking, there is no chance that gatekeeping will be done appropriately.

The gatekeeper needs to coordinate and balance reasons belonging to different rationalities, and that may, in some situations, be difficult to integrate into the already heterogeneous practice of medicine. Physicians are used to finding themselves in situations where different rationalities prevail and in which a balance needs to be struck.³⁴ It is the weighing of budgetary and justice considerations, quite apart from medical ones that is relatively new. The only way to prevent the different obligations of gatekeepers from becoming conflicting loyalties is by providing them with the necessary resources and the training to perform their "balancing act".

Acknowledgements

I wish to thank the two referees of the journal and Marjan Veldhuis, Bert Schadé and Theo van Willigenburg for their useful comments on earlier versions of the paper. This paper was written within the framework of the Incentive Programme Ethics and Policies, which is supported by the Netherlands Organisation for Scientific Research (NWO).

Dick L Willems, MD, PhD is a General Practitioner and a Senior Researcher in Ethics at the Department of General Practice, Division of Public Health, Academic Medical Centre and University of Amsterdam, Amsterdam, the Netherlands.

References

- Gérvás J, Perez Fernandez M, Starfield BH. Primary care, financing and gatekeeping in western Europe. *Family Practice* 1994;11:307-17.
- World Health Organization. *Draft charter for general practice/family medicine in Europe. Report on a WHO meeting. EUR/HFA target 28.* Copenhagen: Regional WHO Office, 1998.
- Glasgow N. The gatekeeper controversy: why it exists and how it can be resolved. *New Zealand Medical Journal* 1996;109:168-70.
- Anonymous. Storming the gates (the gatekeeper). *Gastroenterology* 1997;112:1784.
- Herd B, Herd A, Mathers N. The wizard and the gatekeeper: of castles and contracts. *British Medical Journal* 1995;310:1042-4.
- Mouradian WE. Who decides? Patients, parents, or gatekeepers: pediatric decisions in the craniofacial setting. *Cleft Palate-Craniofacial Journal* 1995;32:510-4.
- Siegel-Itzkovich J. Israel moves to GP gatekeeper system. *British Medical Journal* 1998;317:432.
- Halm E, Causino N, Blumenthal D. Is gatekeeping better than traditional care? A survey of physicians' attitudes. *Journal of the American Medical Association* 1997;278:1677-81.
- Angell M. The doctor as double agent. *Kennedy Institute Ethics Journal* 1994;5:279-86.
- Sprung CL, Eidelman LA, Steinberg A. Is the physician's duty to the individual patient or to society? *Critical Care Medicine* 1995;23:618-20.
- Kassirer J. Managing care—should we adopt a new ethic? *New England Journal of Medicine* 1998;339:397-8.
- Toulmin S. Divided loyalties and ambiguous relationships. *Social Science and Medicine* 1986;23:783-7.
- Kao AC, Green DC, Zaslavsky AM, Koplan JP, Cleary PD. The relationship between method of physician payment and patient trust. *Journal of the American Medical Association* 1998;280:1708-14.
- Caplan AL. Can money and morality mix in medicine? *Academic Emergency Medicine* 1994;1:73-81.
- Mechanic D, Schlesinger M. The impact of managed care on patients' trust in medical care and their physicians. *Journal of the American Medical Association* 1996;275:1693-7.
- Sulmasy DP. Physicians, cost control, and ethics. *Annals of Internal Medicine* 1992;116:920-6.
- Chiu RC. Patient advocate versus societal gatekeeper and the future of high-tech medicine. *Pacing & Clinical Electrophysiology* 1994;17:135-7.
- Sulmasy DP. Physicians, cost control and ethics. *Annals of Internal Medicine* 1992;116:920-6.
- Bernat JL. Quality of neurological care. Balancing cost control and ethics. *Archives of Neurology* 1997;54:1341-5.
- Bodenheimer T, Lo B, Casalino L. Primary care physicians should be coordinators, not gatekeepers. *Journal of the American Medical Association* 1999;281:2045-9.
- Council on Ethical and Judicial Affairs A. Ethical issues in managed care. *Journal of the American Medical Association* 1995;273:330-5.
- Walzer M. *Spheres of justice.* Oxford: Blackwell, 1993.
- Mathers N, Usherwood T. The gatekeeper and the wizard revisited. *British Medical Journal* 1992;304:969-71.
- Mathers N, Hodgkin P. The gatekeeper and the wizard: a fairy tale. *British Medical Journal* 1989;298:172-4.
- Pellegrino E. Rationing health care: the ethics of medical gatekeeping. *Journal of Contemporary Health Law and Policy* 1986;2:45.
- Gillon R. Ethics, economics, and general practice. In: Mooney G, McGuire A, eds. *Medical ethics and economics in health care.* Oxford: Oxford University Press, 1988: 114-34.
- Callahan D. Necessity, futility, and the good society. *Journal of the American Geriatrics Society* 1994;42:866-7.
- Ubel P, Arnold R. The unbearable rightness of bedside rationing. *Archives of Internal Medicine* 1995;155:1837-42.
- Morreim EH. *Balancing act. The new medical ethics of medicine's new economics.* Dordrecht: Kluwer Academic Publishers, 1991.
- Thévenot L, Boltanski L. *De la justification. Les économies de la grandeur.* Paris: Gallimard, 1991.
- Tomlin Z, Humphrey C, Rogers S. General practitioners' perceptions of effective health care. *British Medical Journal* 1999;318:1532-5.
- Fleck LM. Just caring: health reform and health care rationing. *Journal of Medicine and Philosophy* 1994;19:435-43.
- Burg W, van Willigenburg T. Introduction. *Reflective equilibrium. Essays in honour of Robert Heeger.* Dordrecht/Boston/London: Kluwer Academic Publishers, 1998.
- Gillon R. Ethnography, medical practice and moral reflective equilibrium. *Journal of Medical Ethics* 1996;22:259-60.
- Daniels N. *Justice and justification. Reflective equilibrium in theory and practice.* Cambridge: Cambridge University Press, 1996.
- Nagel T. The fragmentation of value. In: *Mortal questions.* Cambridge: Cambridge University Press, 1979: 128-41.
- Beauchamp T, Childress J. *Principles of biomedical ethics.* New York: Oxford University Press, 1994.