A defence of medical paternalism: maximising patients' autonomy

Mark S Komrad  Duke University School of Medicine, Durham, North Carolina, USA

Editor's note

All illness represents a state of diminished autonomy and therefore the doctor-patient relationship necessarily and justifiably involves a degree of medical paternalism argues the author, an American medical student. In a broad-ranging paper he discusses the concepts of autonomy and paternalism in the context of the doctor-patient relationship. Given the necessary diminution of autonomy which illness inflicts, a limited form of medical paternalism, aimed at restoring or maximising the patient's autonomy is entirely acceptable, and indeed fundamental to the relationship he argues. However, the exercise of this paternalism should be flexible and related to the current 'level of autonomy' of the patient himself. An editorial in this issue comments briefly on this paper.

The physician-patient relationship is surely characterised by certain types of inequalities among the two participants. Perhaps the least disputed is the knowledge gap that separates patient and doctor. It motivates the former to seek medical attention in the first place and underlies, in part, the professional authority of the latter. This particular asymmetry has historically been used to justify medical paternalism at the alleged expense of the patient's autonomy. Paternalism has been one of the traditional characteristics of the therapeutic relationship in medicine that distinguishes it from a mere contractual interaction of coequals, perfect mutuality, and simple negotiated claims. However, it is precisely this feature of the doctor-patient relationship that has suffered the harshest criticism lately with the advent of medical consumerism, self-help, the patients' rights movement, and the reevaluation of professional authority in general (1). Many critics have inveighed against paternalism in medicine, equating it with presumptuousness and condescension. Some have advocated the complete extirpation of paternalism from this setting and would bestow perfect autonomy and complete responsibility upon the patient. They have called attention to the 'generalisation of expertise' whereby a physician's technical expertise has been falsely confused with moral expertise (2). Others have sought to combine liberal proportions of autonomy with some paternalism in a less radical reformation. It would indeed be timely to examine the notions of paternalism and autonomy and come to some conclusion about their appropriate interaction in the context of the doctor-patient relationship.

The principle of autonomy

Perhaps the most frequently cited discussions of autonomy are those of Immanuel Kant and John Stuart Mill. Kant's deontological concept, known as 'autonomy of will', is by no means contradictory to Mill's utilitarian 'autonomy of action' but complementary. Kant's notion of autonomy is focused on the rational human will. In the *Groundwork of the Metaphysic of Morals* Kant explains that free will inherent in human beings is the true province of autonomy and as such, autonomy exists prior to action. The autonomous will is both self-governing and self-legislating. It is 'not merely subject to the law, but... must be considered as also making the law for itself' (3). Action, which may flow from the exercise of the autonomous will, is a separate issue from autonomy altogether. One of the special properties of the self-legisitating, autonomous will is that the laws it makes for itself turn out to be universal laws. Kant supplements his definition of autonomy with the mandate that we are compelled to be autonomous. Persons have a duty to be autonomous, especially since autonomy is the basis of all other moral behaviour. It is a 'categorical imperative' not only to follow universal law but to follow a universal law which we ourselves make as moral agents' (4).

For Mill, the principle of autonomy does not arise in the prior will to act but from subsequent action itself. Mill uses the word 'liberty' when discussing autonomy in order to connote freedoms and restraints of action in the context of society. Autonomy is defined by exclusion as freedom of action only in so far as others are not harmed. 'The principle [of autonomy] requires liberty of tasks and pursuits, of forming the plan of our life to suit our own character; of doing as we like,
subject to such consequences as may follow: without impediment from our fellow creatures, so long as what we do does not harm them' (5). Mill's principle of autonomy then is a functional one. He begins with the assumption that autonomy implies that any action which is self-serving is permissible within limits. Those limits are derived by an arm-chair projection of the consequences of hypothetically committed actions. This is followed by a moral calculus: 'conduct from which it is desired to deter [a person] must be calculated to produce evil to someone else' (p 135, emphasis mine). Autonomy is the liberty to execute any action within the bounds that are mapped out in this way.

As it is for Kant, autonomy is a moral imperative for Mill. We are obligated both to fulfil our own potential for autonomy and to preserve the autonomy of others. Mills's autonomous agent may not act in ways that would diminish his own autonomy, as for example by selling himself into slavery. Autonomy is to be preserved at all cost. 'The principle of freedom cannot require that [a person] should be free not to be free. It is not freedom to be allowed to alienate his freedom' (p 236).

In summary, Mill's principle of autonomy is connected with overt actions based on self-interest while Kant's centres on the prior will to act. The former describes the external ordering of autonomy while the latter examines the internal. These two views are clearly complementary. Combining the two we may see autonomy as a self-determined organisation of will according to a *priori* universal laws and also a liberty to pursue self-regarded actions in so far as they do not harm others.

Both deliberation and action are salient features of the concept of autonomy outlined by Beauchamp and Childress: 'The autonomous person is one who not only deliberates about and chooses such plans, but who is *capable* of acting on the basis of such deliberations' (6). The element of capability in this definition raises an important question. Is autonomy a 'natural right'? Are there any who are legitimately excluded from being autonomous on the basis of incapacity and is this proper grounds for restricting autonomy? Many philosophers have recognised that the principle of autonomy is not absolute and that there are those who cannot and should not enjoy a full measure of it. These thinkers identify certain qualifications for full autonomy such as 'capability' and 'maturity' (7, 8). Mill attaches a reminder to his description of autonomy that, 'it is perhaps hardly necessary to say that this doctrine is meant to apply only to human beings in the *maturity* of their faculties'. By 'maturity' Mill means, 'capable of being improved by free and equal discussion' (pp. 135–136, emphasis mine). Beauchamp and Childress contend that 'some persons are not in a position to act in a sufficiently autonomous manner, perhaps because they are immature or incapacitated. Infants and irrationally suicidal individuals are typical examples' (9). Maturity however is not a simple state but varies by degree in the process of 'maturation'. If it is legitimate to link autonomy and maturation (capability), as has been suggested, then one can conceive of a 'degree of autonomy' that fluctuates with time and situation as does capability. This point will become important later when considering the place of autonomy in the therapeutic relationship.

### The concept of paternalism

The question of who are properly autonomous and the degree to which they are so, conveniently invites a consideration of paternalism. When autonomy recedes paternalism advances and vice-versa. Paternalism cares for an individual's interest in place of autonomy, either by force or by necessity. Thus, paternalism and autonomy are two inversely varying parameters along a spectrum of independence. However, they are by no means entirely contrapositives. The object of both is the good of the same moral agent. From the perspective of autonomy this good appears as self-interest while it is conceived by the paternalist as a fiduciary interest. Gerald Dworkin's definition highlights this point which is the essence of paternalism: 'By paternalism, I shall understand roughly the interference with a person's liberty of action, justified by reasons referring exclusively to the welfare, good, happiness, needs, interest, or values of the person being coerced' (10).

Dworkin's definition relies on Mill's view of autonomy as 'liberty of action' and unfortunately sees paternalism as always coercive. Others have perceived the themes of superiority, domination, oppression, and dogmatism in paternalism (11). When characterised in this way it is more properly called authoritarianism. Paternalism is not necessarily coercive behaviour; there is another side to it that connotes the concern, care and self-sacrifice of the paternalistic agent. Nor does the notion of coercion fully characterise paternalism, which may manipulate thought and information as well as action. Dworkin's interpretation does not give credit to that component of autonomy prior to action in which Kant is interested.

Gert and Culver offer a more temperate view of paternalism. They define paternalistic behaviour as follows (12):

A is acting paternalistically toward S if and only if A's behaviour (correctly) indicates that A believes that:
1. his action is for S's good;
2. he is qualified to act on S's behalf;
3. his action involves violating a moral rule (or will require him to do so) with regard to S;
4. S's good justifies him in acting on S's behalf independently of S's past, present, or immediately forthcoming (free, informed) consent, and
5. S believes (perhaps falsely) that he (S) generally knows what is for his own good.

This definition is compelling because it is clear, general enough to be useful, and has some novel properties.
According to this view, a paternalistic act does not necessarily violate liberty of action, as Dworkin held, but 'moral rules' which the authors relate to 'rights'. Only a small subset of these moral rules concern liberty of action (13). Like Dworkin's, this definition also emphasises the purely fiduciary motives inherent in paternalism. In fact, feature 1 is the "sine qua non" of paternalism. 'What makes A's action toward S paternalistic is never the good of anyone other than S himself' (14).

Feature 4 is also a crucial one and is certainly unique. An act is not paternalistic if S gives consent. 'Past, present or immediately forthcoming consent removes A's act from the class of paternalistic acts' (15). Nevertheless, the paternalist must have a reasonable expectation of S's eventual consent (16). This suggests that it is impossible to grant a request for paternalism since any response to such a request is, by definition, not paternalistic but remunerative; 'solicited paternalism' is a contradiction in terms. If this point is accepted then an interesting result follows. It has been popular to consider cases where patients voluntarily relinquish their autonomy in the name of 'doctor-knows-best' (and who have not been coerced to do so). Some thinkers have attempted to argue that this is a rare example of 'morally justified' paternalism (17). Holding to Gert and Culver's terminology, this is a moot argument and the issue must be entirely reclassified in some other area than the ethical debate over paternalism. For a physician to accept such an offer would not be paternalistic. However, when a patient voluntarily surrenders his autonomy one would still like to have some constraints on the physician's behaviour in order to ensure the sentiment of trust. Although features 3 or 4 may not be fulfilled in this case, a doctor may still be enjoined to observe 1, 2, and 5. Gert and Culver denote such behaviour, which would be paternalistic except for the presence of consent, as 'paternal behaviour'. Paternal behaviour does not violate moral rules and occurs in response to a request or consent that is not elicited by coercion.

In this context, Mill might consider it morally reprehensible for a patient completely to waive his or her autonomy. Oddly, the paternalistic act would be for the physician to refuse this offer and force the patient to retain his autonomy, perhaps even against his will. Relinquishing autonomy to a physician is analogous to selling oneself into slavery, except that it may not be quite so permanent. Mill claims that a State is justified in paternalistically prohibiting a man from selling himself into slavery. We are not free, he argued, to moderate our freedom nor deliberately to abandon our autonomy. Kant's 'categorical imperative' would also forbid even the voluntary and informed renunciation of autonomy. Ironically then their arguments lead to the conclusion that one may paternalistically deter a patient from inviting paternalism at the expense of his autonomy. By the above distinction, to refuse such a request would be paternalistic, to grant it would be paternal.

**Paternalism and autonomy in medical practice**

Having teased out some of the issues contained in the principles of autonomy and paternalism in general, it remains to explore their proper application in the context of the physician-patient relationship. Mark Siegler succinctly reviews the dilemma (18):

The principle of respect for autonomy surely recognises that different autonomous individuals will wish to be treated in different ways by the health professional. . . . The critical question . . . [is how] morally conscientious physicians and patients . . . determine where on a spectrum of paternalism/consumerism or dependence/independence their professional relationship will and ought to stabilise.

Some of the early codes of the medical profession seemed to promulgate absolute paternalism as an appropriate professional etiquette (19). The priest-like status of doctors historically encouraged paternalism to which patients readily acquiesced. However, in actual practice, there is evidence that physicians usually adjusted this paternalism to the context and the patient; they were more responsive than dogmatic (20). Recently, and especially in the United States, the 'crises of authority', of which medicine has been a part, have been accompanied by a virtual apotheosis of autonomy and a profanation of paternalism. Patient-consumers have campaigned for 'respect for autonomy' and have cast the issue in the language of human rights. Many reformers have called upon contract ethics as a model for the therapeutic relationship, eschewing even minimal paternalism (21).

In general, many physicians have tried to rise to the occasion and curtail paternalistic tendencies, nourished in the course of medical education and socialisation (22). The law too has helped to bolster incentive in this direction. Yet it is somewhat confusing for a doctor when, on the one hand, society insists on the abolition of medical paternalism, while on the other hand many individual patients still expect, hope for, and even urge (in both subtle and outright ways) the doctor to be paternalistic. Physicians have been faced with the option of indulging these expectations and receiving accolades from the patient but condemnation from society.

Consider this typical example. A physician discovers a 1.5 cm breast cancer in a 30-year-old woman without swollen axillary lymph nodes. He plans a total mastectomy instead of local excision of the mass followed by radiation therapy, which is controversially of equivalent efficacy (23). The patient requests her doctor to do 'whatever he thinks best'. Can the physician accept this invitation? Would an attempt to offer the patient a justification for mastectomy be superfluous in so far as the patient has already decided that the doctor's judgment is unconditionally acceptable? Should the doctor aggressively force autonomy back on the patient, insisting that all arguments for and against mastectomy...
be presented for the patient's scrutiny? Suppose the patient insists on flipping a coin to decide between therapeutic alternatives. The principle of absolute autonomy would not restrict a patient, who otherwise seems to have an idea of her own good (Gert and Culver's feature 5), from using a non-rational technique to make a choice. When autonomy becomes capriciousness is it still sacred? Is it presumptuous to denote when it has become so?

What is more confusing yet is that patients will sometimes give mixed messages. A patient who may at first insist on a strict non-paternalistic attitude from his physician may later castigate the doctor for not having been more paternalistic, when autonomously chosen decisions lead to disaster. Consider the 50-year-old man with angina who announces to his physician that he is taking up jogging by joining a club. After an initial evaluation the reluctant doctor signs the medical form but urges caution. The patient jogs without incident for three months and indeed, his cardiac parameters improve; until one day on the track he has a severe bout of angina and is brought to the emergency room with a myocardial infarction. When he recovers he strongly berates his doctor for not forbidding him to run at the very first, and for not using paternalistic blackmail by refusing to prescribe him medication unless he gave up running. Patients often want it both ways depending on their convenience and not on moral principle. There is thus a variation in the demand for autonomy from patient to patient, and from patient to society. A single patient may vary his own needs and demands in the course of a therapeutic encounter. Therefore, a formula for medical paternalism cannot be dogmatic and unconditional. This point merits closer attention.

Preserving some paternalism

Recall that many philosophers identify a category of persons who are legitimate candidates for paternalism: people who 'do not culture reason' (Kant), the immature (Mill), the inherently 'non-autonomos' (Beauchamp and Childress), etc. This suggests that there are some human conditions in which people are not capable of enjoying a full measure of autonomy and that paternalism should protect their interests where autonomy is wanting. I earlier proposed that capability should be a determinant of 'degree of autonomy'. Individuals with impaired capacities suffer diminished autonomy. Paternalism is actually a response to this incapacity and not a negation of rights. Surely the human condition is protean and people occasionally experience diminished or imperfect autonomy, often only temporarily. Autonomy is neither permanent nor immutable, but is a dynamic state liable to perturbation. Some common examples where autonomy is lessered are imprisonment, pregnancy, marriage, and political office. However, the most striking examples are illness and disease: an accident causing quadriplegia, the development of schizophrenia, stroke, tuberculosis. In fact, I maintain that all illness represents a state of diminished autonomy. The ill are dependent on others such as physicians, if not for outright therapeutic ministrations then for their expert legitimation of their illness.

Talcott Parsons describes the dependency and vulnerability of the sick role that make it an involuntary state of diminished autonomy: 'By definition of the sick role, the sick person is helpless and therefore in need of help. . . . He is not only generally not in a position to do what needs to be done but he does not know what needs to be done or how to do it' (24). However, there are clearly stronger reasons to describe illness as a state of diminished autonomy than that the sick are partially ignorant and thus dependent on their physicians. Otherwise, ignorant owners of broken cars would not be autonomous; and sick (but knowledgeable) physicians would be fully autonomous. In fact, illness is qualitatively different from owning a broken car - profoundly so. It is an existential condition that gently or harshly impinges on the human soul. Leon Kass waxes lyrical in describing the physician who:

must tend particular, necessitous human beings who, in addition to their symptoms, suffer from self-concern and often fear and shame - weakness and vulnerability, neediness and dependence, loss of self-esteem, and the fragility of all that matters to them . . . medicine deals - sometimes explicitly - with the fact of human embodiment, that is, with our strange and mysterious being comprising a grown-togetherness of soul and body (25).

Pellegrino calls illness 'an ontological assault aggra-vated by the loss of freedoms we identify as peculiarly human' (26). Thus, illness is qualitatively and uniquely different from other more mundane situations of diminished autonomy which are due to partial ignorance alone.

Illness is an attenuation of autonomy in both Kantian and Millian terms. Physical incapacity mitigates the liberty of action and thus diminishes 'autonomy of action', in Mill's sense. Mental or even physical illness can interfere with reason and thus deprive one of the faculty that is crucial for Kant's 'autonomy of will'. Parsons observes that even in purely physical illness, 'the situation of the patient is such as to make a high level of rationality of judgment difficult' (27).

When seen in this way - as a special kind of diminished autonomy - the sick role naturally invites the physician to behave paternalistically if not frankly paternalistically, to fill the void left as autonomy diminishes. Incidentally, one would not want the physician to behave in any other way since paternalism is the only type of response that properly puts the patients' good above all other considerations. According to this view, some paternalism is not only justified but is required in all therapeutic relationships due to the nature of illness and the sick role. Paternalism is not always incompatible with the principle of autonomy and, in fact, pater-
Szasz and Hollender developed a dynamic model of the therapeutic relationship based on the principle that it is 'a process, in that the patient may change not only in terms of his symptoms but also in the way in which he wishes to relate to his doctor' (31). They outline three reference points along the continuous spectrum of dependence-independence, paternalism-autonomy.

One extreme is **activity-passivity** where the patient is inert and the doctor does things to him without consent or dissent. The physician is absolutely paternalistic here. The prototype of this state is the parent-infant relationship. At the opposite extreme is complete **mutual participation**. In this condition, the doctor and patient are equal, mutually independent, and are anxious to satisfy one another. This is following the contract ethics model. This relationship is appropriate when the patient's autonomy is truly just sub-maximal (it is never maximal as long as he is in the sick role). The doctor 'helps the patient to help himself'. The prototype is the relationship between two independent adults. Between these two extremes is what the authors call **guidance-co-operation**. Here, the patient is far more of a participant than in the first model but is active only as a co-operator and can only partially exercise his judgment. The analogue of this relationship is that of parent and adolescent.

These three states represent only two extremes and a midpoint. The therapeutic relationship should be able to assume any combination of paternalism and autonomy along this continuous scale. Moreover, modulations along the scale should be easy and dictated by negotiation and situation. Some clinical examples where activity-passivity is the rule are anaesthesia, acute trauma and coma. Guidance-co-operation is suited to acute infection, post-operative care, etc. Mutual participation is appropriate for psychoanalysis, chronic illness and rehabilitation. The continuously shifting ratio of paternalism and autonomy is illustrated by a diabetic who enters the emergency room in ketoacidotic coma and is eventually discharged on a responsible diet and insulin regimen. This progression traces the therapeutic relationship from maximal to minimal paternalism. Generally, the most important determinant of where along this spectrum the relationship will stabilise is the degree to which the patient's autonomy is diminished at any one time and must be restored by compensatory limited paternalism.

**Comparison with other approaches**

The scheme which has been proposed here uses the maximisation of autonomy as the touchstone for evaluating paternalism. It helps to reconcile the seeming disparity between autonomy and paternalism which many think cannot and should not co-exist in the therapeutic relationship. This approach is preferable to other techniques of reconciliation such as the 'reasonable man' standard (32), cost-benefit analysis (33), and the moral calculus that pits the evils of inter-

---

**How much paternalism is appropriate?**

Having established that it would be undesirable to dispose entirely of paternalism in medicine, and having examined what type of paternalism in particular should be included, it remains to consider the quantitative problem. That is, how much of this limited paternalism is called for. Clearly, rigid and unconditional formulas are not useful for they neglect the variables of time and situation that are constantly changing boundary conditions for any formula (29). It is more realistic to suggest that the dispensation of limited paternalism be tuned to the peculiarities of context. Just as one can speak of a 'degree of autonomy' so is there a 'degree of paternalism' which the physician must gauge on the basis of the patient's condition (how compromised is his autonomy?), the patient's desires, and the purely technical constraints of the moment (eg unconsciousness). Constant feedback between patient and doctor should enable the balance of paternalism and autonomy to be continuously updated and fine-tuned. This spirit of kinetic reciprocity whereby paternalism is moulded to the situation, is central to Mark Siegler's concept of 'the physician-patient accommodation, [which] is not a permanent, stable and unchanging relationship between physician and patient . . . it is a dynamic model and is always in flux' (30).
ferring with liberty against the evils spared by such interference.

The 'reasonable man' standard suffers from moral relativism that is inherently unsatisfying. It does not acknowledge that a person who has assumed the sick role is not an average, reasonable man. The cost-benefit technique can include definitions of cost, harm and benefit that are so broad as to be useless. Finally, those who utilise a moral calculus to weigh injustices suffered by the patient are in danger of presumptuousness. Thus, one of the views of paternalism is that the patient's degree of incapacity, his immediate cost, harm, and benefit are viewed as a possible justification for paternalism. One's degree of incapacity may be compromised. Both philosophers insist that autonomy is not so much a right as it is a duty to pursue according to one's capabilities.

Patrialism is acting in another's interest in the absence of his or her immediate consent, although with the expectation of eventual consent. It is first cousin to autonomy since both are related to the same good of the same person. Thus, paternalism and autonomy are reciprocal. Where autonomy falters paternalism supports.

Autonomy is not universal and there are persons who have varying capacities to behave autonomously. There are 'degrees of autonomy' based on capability that legitimately require compensatory 'degrees of paternalism'. One's degree of autonomy is not fixed but fluctuates in the course of human affairs. Illness can be viewed as a state of diminished autonomy and this is in fact one of the important features of the sick role in our culture. Thus, the sick constitute a group with less autonomy than when they were healthy and, as such, they require some element of paternalistic treatment. However, this is a very limited type of paternalism, designed solely to maximise the patient's autonomy. This is the only legitimate interest of medical paternalism. The patient's incapacity for autonomy as a result of his condition is the major factor determining where on the spectrum of paternalism-autonomy the thera-

peutic relationship must operate. Paternalism is a response to incapacity, not a negation of rights. A continuous update of the patient's autonomy-status is required to modulate the doctor's paternalism.

It is important to recognise that the doctor-patient relationship is a dynamic process. It is a journey from limited paternalism to maximal autonomy which is its telos, or ultimate purpose. As the patient's capacity for autonomy increases, so the physician's paternalism which nurtures that autonomy decreases. The recovery of a patient from a diabetic coma can be used to illustrate this process.

The relationship between patient and doctor is unique among the professions, not because of the knowledge gap but because of the special dimensions of anguish in illness (35). Confronted with this reality, it is difficult to deny the patient's need, however slight, for a paternalism which is not a challenge to his autonomy but its champion.

Acknowledgments

I am most grateful to the staff at the Hastings Center for Society, Ethics, and the Life Sciences for their valuable discussions concerning the issues raised in this paper, and for the use of their splendid resources.

This work was supported by the National Institutes of Mental Health, Medical Student Psychiatric Education Grant, MH15204; The Josiah Charles Trent Memorial Foundation Grant, 383-73578; and a student internship grant from The Hastings Center for Society, Ethics, and the Life Sciences.

References and notes

(1) Haug E, Lavin B. Public challenge of physician authority. Medical care 1979; 17: 844-858.
(4) See reference (3): 34.
(9) See reference (6): 60.
References and notes


(2) This redefinition of death is advocated by Veatch R. The whole-brain-orientated concept of death: an outmoded philosophical formulation. In: Beauchamp and Walters, see reference (1): 267-276.

(3) This alternative to redefining death is proposed by Dworkin R. Death in context. Indiana law review 1973, Summer; 48: 623-646. The above examples of how the law could be reformulated come from him.

(4) These further possibilities are suggested by Jonas H. In: Beauchamp and Walters, see reference (1): 265.


(6) A full-scale refutation of medical paternalism lies outside the scope of this essay, but I think the essential weaknesses of that position have been ably exposed by Beauchamp T L. Paternalism and biobehavioural control. In: Beauchamp and Walters, see reference (1): 522-529; and Buchanan A. Medical paternalism. Philosophy and public affairs 1978, Summer; 7, No 4: 370-390.