Rationing in health systems: A critical review

Iman Keliddar1, Ali Mohammad Mosadeghrad2*, Mehdi Jafari –Sirizi3

Received: 20 Jan 2017      Published: 27 Aug 2017

Abstract
Background: It is difficult to provide health care services to all those in need of such services due to limited resources and unlimited demands. Thus, priority setting and rationing have to be applied. This study aimed at critically examining the concept of rationing in health sector and identifying its purposes, influencing factors, mechanisms, and outcomes.

Methods: The critical interpretive synthesis methodology was used in this study. PubMed, Cochrane, and Proquest databases were searched using the related key words to find related documents published between 1970 and 2015. In total, 161 published reports were reviewed and included in the study. Thematic content analysis was applied for data analysis.

Results: Health services rationing means restricting the access of some people to useful or potentially useful health services due to budgetary limitation. The inherent features of the health market and health services, limited resources, and unlimited needs necessitate health services rationing. Rationing can be applied in 4 levels: health care policy-makers, health care managers, health care providers, and patients. Health care rationing can be accomplished through fixed budget, benefit package, payment mechanisms, queuing, co-payments, and deductibles.

Conclusion: This paper enriched our understanding of health services rationing and its mechanisms at various levels and contributed to the literature by broadly conceptualizing health services rationing.

Keywords: Health system, Health Care Rationing, Health resources, Patient selection, Critical review

Introduction
The World Health Organization defined health as a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity (1). Accordingly, governments should establish robust health systems to promote the health status of their nations. Every human being has the right to have access to the highest attainable standards of health without distinction of gender, race, religion, belief, income, or social class (1). Providing high quality, accessible, and affordable health services should be the main goal of health care organizations (2).

Health care organizations are facing several challenges such as unlimited demands, increasing costs, and resource shortage (3-4). During the last 15 years, the rate of health expenditure growth exceeded the rate of economic growth (5). However, money, especially in developing countries, is disproportionately spent on those health services that have a low impact on people’s health, and mainly benefit the rich (6). Inefficiencies in the delivery of health care services, variable providers’ payment systems, supplier induced demand, and inappropriate use of expensive technologies are the main causes of escalating health care costs (7). For instance, in some OECD countries, and the US, new advanced technologies with less marginal effectiveness were sometimes used instead of alternative and less expensive existing procedures (8). Redirecting even a fraction of that money could expand useful health care coverage, enhance quality of care, and improve patient satisfaction.

It is highly important to get the best value for money due to the shortage of health care resources. Thus, health
Rationing in health systems

Table 2: Criteria for quality appraisal of the papers

- Are the research goals and objectives clearly specified?
- Is the research design clearly specified and is it suitable for achieving research goals?
- Is the research process clearly explained?
- Are enough data displayed to support research interpretations and conclusions?
- Is the analysis method appropriate and adequately explained?

Source: Dixon-Woods, et al. 2006

Health care rationing literature is large, complex, and diverse. Thus, we decided to use critical interpretive synthesis (CIS), instead of conventional systematic review methodology, to achieve the aims of this study. Critical interpretive synthesis enables researchers to critically synthesize a diverse body of evidences and studies in a given field and generate theories. While conventional systematic review aims to test theories through searching for, appraising, and synthesizing the findings of the primary studies (aggregative syntheses) in a fixed procedure of predefined sequence, critical interpretive synthesis aims to generate a theory by including many different forms of evidence (Interpretive syntheses) through iterative, dynamic, interactive, and recursive processes of question formulation, searching for, data extraction, critique, and synthesis (11).

In this review, which was conducted during the summer of 2015, PubMed, Cochrane, and Proquest databases were searched to find published English literature in health services rationing from 1970 to 2015. Keywords were rationing health *, priority setting health *, setting priority* and services rationing from 1970 to 2015. Keywords were rationing health services were described. In the first and second sections, the concept of rationing and its necessity in health sector were discussed. The third part was devoted to the trends of rationing in the health sector. Then, the levels of applying health care rationing were examined. Finally, the mechanisms for rationing health services were described.

What is health services rationing?

There is no universal definition of the word ‘rationing’. Fourteen definitions of rationing were elicited from the literature (Table 2). The word ‘rationing’ has derived from the Latin ‘ration’, meaning ‘to apply limitation in usage’ (13). In economics, rationing refers to controlling the distribution of scarce resources and services among a population (14). Oxford Dictionary defines rationing as ‘allowing each person to have only a fixed amount of a commodity’ (15).

Ubel defines rationing as “any implicit or explicit mechanisms that allow people to go without beneficial services” (16). Russell (2002) defined rationing as “someone or some institution’s deliberate decision to distribute a scarce good among competing persons” (17). According to Brown (1991), rationing is the “deliberate and systematic withholding of beneficial goods or services from some elements of the population on the grounds that the society cannot afford to extend them” (18).

Scarcity, resource constraint, and exclusion are 3 important notions that constitute the definition of health care rationing. Therefore, we define health services rationing as “restricting the access of some people to useful or potentially useful health services due to budgetary limitation”.

Why health services must be rationed?

Products or services should have 3 features to be rationed. These features are scarcity, value, and controllability (17). All these features are inherent in health care services, which make rationing a necessity. First, health care

Table 1.

<table>
<thead>
<tr>
<th>What is health services rationing?</th>
</tr>
</thead>
</table>
| There is no universal definition of the word ‘rationing’. Fourteen definitions of rationing were elicited from the literature (Table 2). The word ‘rationing’ has derived from the Latin ‘ration’, meaning ‘to apply limitation in usage’ (13). In economics, rationing refers to controlling the distribution of scarce resources and services among a population (14). Oxford Dictionary defines rationing as ‘allowing each person to have only a fixed amount of a commodity’ (15). Ubel defines rationing as “any implicit or explicit mechanisms that allow people to go without beneficial services” (16). Russell (2002) defined rationing as “someone or some institution’s deliberate decision to distribute a scarce good among competing persons” (17). According to Brown (1991), rationing is the “deliberate and systematic withholding of beneficial goods or services from some elements of the population on the grounds that the society cannot afford to extend them” (18). Scarcity, resource constraint, and exclusion are 3 important notions that constitute the definition of health care rationing. Therefore, we define health services rationing as “restricting the access of some people to useful or potentially useful health services due to budgetary limitation”. Why health services must be rationed? Products or services should have 3 features to be rationed. These features are scarcity, value, and controllability (17). All these features are inherent in health care services, which make rationing a necessity. First, health care
is a scarce commodity. The resource shortage for health services delivery is especially noticeable in the developing countries, where a smaller share of resources is allocated to the health sector (31). The problem would be worse during the economic recession. For instance, health spending fell in half of the EU countries between 2009 and 2012, following the economic crisis. As a result, health expenditure declined by 0.6% annually, compared with an annual growth of 4.7% between 2000 and 2009 (32). In addition, the human resources needed for health care organizations are decreasing due to the aging societies and a rise in the demand for health care services in Europe (33).

Second, health is a valuable good and has the value of life principle; and life itself is inseparable from health. Without some degree of healthy functionality, the living whole would not exist. Thus, people may have the willingness to give their assets away in return of getting their health back. Thus, access to adequate health care is considered as a fundamental human right in many countries. Finally, the health services are controllable and can be refused to provide to some people and reserved to meet immediate demands. Therefore, health service is one of the cases which demands rationing.

The goal of rationing is to supply rational, equitable, and cost-effective health services while reducing expenditure (34). Moreover, the inherent characteristics of the health market also make it necessary to apply rationing (Table 3). Characteristics such as information asymmetry and heterogeneity make the use of rationing inevitable. The information asymmetry between patients and physicians may increase supplier induced demand (35). On the other hand, health care professionals deliver the services to patients differently (36, 37). Variations in medical practices raise questions about the equity, quality, and efficiency of health care services (38). Limited health care resources suggest applying rationing mechanisms such as evidence-based guidelines and protocols to deliver just effective care to patients. As a result, induced demand and medical practice variations can be reduced.

**What is the trend of health services rationing?**

The notion of rationing has evolved gradually from an “implicit” and “hidden nonsystematic” to an “explicit and open systematic” mode (39). Implicit rationing relies on hidden norms and rules mainly defined by health care providers such as physicians. In contrast, explicit rationing is based on clearly defined indicators such as patient age, gender, financial status, and clinical condition.

Traditionally, an individual or groups of doctors decide who and when gets treated in an uncoordinated way (40). The public is not involved in the rationing decisions in a “hidden nonsystematic approach. Hence, medical doctors within the given budget limit, allocate resources based on clinical priorities and inter-specialty bargaining power. For instance, they may deny dialysis to patients over the age of 55 years according to an unwritten rule due to limited budgets.

In the 1980s, rationing decisions were started to carry out according to systematic well-defined efficiency considerations. The aim was to maximize the amount of health for a population within the given budget without involving the public in the rationing decisions (41). Doctors had a limited power in determining rationing principles and their applications in practice. Using relatively high de-

---

**Table 2. The Definitions of Health Services Rationing**

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaron and Schwartz (1990)</td>
<td>“The deliberate and systematic denial of certain types of services, even when they are known to be beneficial, because they are deemed too expensive.” (p.418) (19)</td>
</tr>
<tr>
<td>Asch and Ubel (1997)</td>
<td>“Not to provide some beneficial health care services, which are simply too expensive” (p.1668) (20).</td>
</tr>
<tr>
<td>Baily (2003)</td>
<td>“To limit the beneficial health care an individual receives by any means – price or non-price, direct or indirect, explicit or implicit” (p.35) (21).</td>
</tr>
<tr>
<td>Bennett and Chanfreau (2005)</td>
<td>“The controlled distribution of scarce goods or services” (p. 542) (22).</td>
</tr>
<tr>
<td>Brown (1991)</td>
<td>“The deliberate, systematic withholding of beneficial goods or services from some elements of the population on the grounds that society cannot afford to extend them.” (p.30) (18)</td>
</tr>
<tr>
<td>Dougherty (1991)</td>
<td>“The denial of services that are potentially beneficial to some people because of limitations on the resources available for health care” (p.3) (23)</td>
</tr>
<tr>
<td>Goldbeck-Wood (1997)</td>
<td>“Withholding a beneficial treatment because of its costs” (p.146) (24)</td>
</tr>
<tr>
<td>Hadorn and Brook (1991)</td>
<td>“The withholding of necessary services and societal toleration of inequitable access (for example, based on ability to pay) to services acknowledged being necessary by reference to necessary care guidelines.” (p.3331) (26)</td>
</tr>
<tr>
<td>Hurst and Danis (2007)</td>
<td>“Any clinical decision to place or to accept a limit on benefits for a patient” (p.248) (27)</td>
</tr>
<tr>
<td>Maynard (1999)</td>
<td>“An individual is deprived of care which is of benefit (in improving health status, or the length and quality of life), which is desired by the patient.” (p.6) (28)</td>
</tr>
<tr>
<td>Ole Frithjof (1999)</td>
<td>“The withholding of potentially beneficial health care through financial or organizational features of the health care system in question.” (p.1426) (29)</td>
</tr>
<tr>
<td>Pickard and Sheaff (1999)</td>
<td>“Restricting access to health care for nonclinical reasons such as cost control” (p.38) (30).</td>
</tr>
<tr>
<td>Ubel (2001)</td>
<td>“Any implicit or explicit mechanism that allows people to go without beneficial services” (p.35) (16).</td>
</tr>
</tbody>
</table>

---

**Table 3. Determinants of rationing health services**

<table>
<thead>
<tr>
<th>Determinants of rationing</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of health services</td>
<td>• Value of health services</td>
</tr>
<tr>
<td></td>
<td>• Scarcity of health services</td>
</tr>
<tr>
<td></td>
<td>• Controllability of health services</td>
</tr>
<tr>
<td>Characteristics of health market</td>
<td>• Information asymmetry</td>
</tr>
<tr>
<td></td>
<td>• Heterogeneity</td>
</tr>
</tbody>
</table>

---

[http://mjiri.iums.ac.ir](http://mjiri.iums.ac.ir)

ductibles and copayments to discourage people from using expensive health services is an example of hidden systematic rationing.

Finally, open systematic rationing has been used since 1990s in some developed countries such as UK (42), where a public recognition was happening. In some countries, a minimum health care package was introduced to ensure people’s access to those services, and health services outside of this package were not guaranteed to the population.

Who does health services rationing?
Health services rationing decisions could be taken by various organizations, groups, and individuals in different levels of a health system. Mechanic (1997) argues that decisions about health services rationing are normally made at 3 levels: health systems, intermediate, and clinical levels (34). At the health care systems level, governments, health authorities, or health insurance companies determine the total health care spending levels, the types of health services to be covered, the extent of technological development, the location of health care facilities, and the extent of patient cost sharing. At the intermediate level, subunits such as hospitals determine the number and mix of various providers, the extent of direct access, schedules, and waiting times for various health care services. Finally, at the clinical level, clinicians decide about treatment priorities considering types of patients, varying conditions, and treatment approaches.

Krizova and Simek (2002) believe that rationing occurs at political and clinical levels (43). Health care politicians, health insurance companies, and hospital executives set up an external economic framework for clinical work. At the clinical level, physicians and medical professionals are responsible for quality of care and carry out the rationing. Klein (1997) suggests that rationing could occur at macro, meso- and micro-levels (44). The macro-level rationing refers to decisions about how much funding should be allocated to health services altogether. Allocation of resources between particular services and localities occur at the meso-level. Finally, rationing at the micro-level deals with decisions on treating individual patients. Coast et al. (1996) propose 4 levels for rationing health services: (a) across whole services, (b) across treatments within services, (c) within treatments, and (d) between individual patients (45).

Therefore, considering the complex nature of health care organizations, rationing could be practiced at 4 levels: patient, provider, manager, and policymaker (Fig. 1). The individual patient’s needs and preferences may lead to self-rationing. A sick person without a health insurance plan makes the decision to seek care by comparing the costs and benefits of receiving the care. S/he might refuse to proceed to get the services if could not afford it or see the costs more than the benefits. Price-based self-rationing leads to underutilization of health care services. The health care team consists of the individual physicians and care providers. Basically, rationing decisions are made according to the clinical interactions between physicians and patients. Despite a commitment to ration health care services openly in theory, clinicians use more implicit methods in practice.

The third level of the health services rationing occurs at the organizational level (e.g. hospital, clinic, and nursing home) that provide infrastructure and needed resources to support health care providers. Managers are responsible for allocating resources among various health care providers and may ration some health services.

Finally, politicians, policymakers and health insurance companies at the macro-level set some rules and regulations about the allocated health care budget, services coverage, and cost sharing mechanisms considering political, economic, social, and technological factors. The comparison between different levels of health services rationing proposed by different scholars and this study is demonstrated in Table 4.

Health care rationing also depends on the financing sources of a health system. While in tax-based health systems, the government authorities could keep the responsibility for health care rationing, in payroll and premium-based health systems, social or private insurance companies deal with the responsibility. Finally, in an out of pocket-based health system, health care institutes’ managers, clinicians, and even patients themselves make the decisions for rationing health care services.

Rationing can be applied in all 4 functions of the health systems including governance, financing, resource generation, and health services provision. At the macro-level, politicians and health care authorities, as the stewards of the national resources, plan and direct rationing health care services to establish the fairest possible health system, to make the best possible use of limited resources,

<table>
<thead>
<tr>
<th>Table 4. Levels of Health Services Rationing in Health Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Klein (1992)</strong></td>
</tr>
<tr>
<td>Macro level</td>
</tr>
<tr>
<td>Meso level</td>
</tr>
<tr>
<td>Micro level</td>
</tr>
</tbody>
</table>

[http://mjiri.iums.ac.ir](http://mjiri.iums.ac.ir)
and to deliver the best health outcomes. The government and Ministry of Health should regulate and balance rationing at meso- and micro-levels to maximize the probability of success in health services rationing. A lack of rationing health care services policy at the political level may lead to uncontrolled medical professionals’ power of decision-making.

How do ration health services?

Health is a right and health services should be equitably accessible to all. Thus, health care services should be rationally distributed to fit the needs of the people who need them. To do so, it is sometimes necessary to limit the access of less needy people to some health care services to provide them to more urgent needs. Several mechanisms could be used for health care rationing, which is classified as supply and demand side mechanisms.

Policymakers and authorities at macro-level can use methods such as budget, benefit package, and payment mechanisms to control the behavior of healthcare managers and providers and restrict providing a broad variety of inappropriate health services. Budget influences the behavior of providers and leads them to reallocate health care resources or even ration some health care services. This method was used in the UK between the 1950s and early 1980s. Budget may even lead to covert rationing. For instance, tight budgets in the UK resulted in the denial of dialysis and some forms of heart surgery to patients over the age of 55 years (5).

Health services coverage is sometimes limited through the specification of a menu of core health care services (minimum health care package) to be made available for the public. Accordingly, certain health care services outside of this package will be unavailable to all patients (46). Practitioners can use clinical guidelines for requesting appropriate effective health care services for patients. Clinical guidelines assist practitioners on which diagnostic tests to order, how to provide surgical, medical, and nursing services and how long patients should stay in the hospital (47). In addition, policymakers use cost-benefit and cost-effectiveness analyses to finalize the benefit package (minimum health care package).

Payment mechanisms as incentives influence the behavior of health care professionals to decrease or increase the number of patients (through capitation), the number of visits (through salary or fee-for-service), the type and quality of services (through fee for services), and the referral of patients to other health care facilities (through salary, fee-for-service) (48).

Some mechanisms such as queuing, copayments, and deductibles can be used at macro-, meso- and micro-levels to limit access to health care services at individual, household, or community levels and control the demands for health care services. Queuing gives some patients a higher priority than others in accessing health care services, which may be explicit by using defined indicators of severity of the patient’s clinical condition, or may be implicit based upon doctor’s referral practices (49). Co-payments and deductibles influence the behavior of the receiver of the health services. High copayments and deductibles discourage but not stop patients from using expensive health care services (5).

Discussion

Health service rationing is a mechanism that could help assure equity (50). Health services rationing is a fundamental issue in health systems’ policymaking. Most studies on health services rationing focus on its specific levels. To the best of our knowledge, this was one of the first studies that offered a theoretical conceptualization of health services rationing through analyzing its whole picture in health systems.

Health services rationing was defined in this study as “to restrict some people’s access to useful or potentially useful health services due to budgetary limitation”. Health care rationing means denying patients the potentially beneficial health care services. Such a definition integrates both clinical and political levels in health rationing. Rationing is about applying price and non-price, direct or indirect, and open or hidden mechanisms to restrict access to the required health care services. These health care services should be provided according to people’s health needs.

Those decisions that limit access to health services based on nonfinancial reasons (e.g., medical reasons) or focus on reducing wastage and administrative costs in health care are not considered as rationing decisions. Only those decisions that cause behavioral changes in different health system actors and consequently limit access to health services can be considered as rationing.

Rationing in health systems is a multi-dimensional field of action. We proposed a model of rationing in health systems based on the system theory (Fig. 2). The health services rationing system is concerned with deciding which goods or services must be restricted and which patients should be given limited access to such services. The principal components of this model are input, process, output,
Rationing in health systems

and feedback. At the first step, a condition occurs in which the demand exceeds the supply of health resources. This gap is made due to a pressure of the environment. The resource shortages and the right to high quality and affordable health services (inputs) feed into the rationing system. Therefore, health system stewardship decides to restrict some people to access health services (process), and this is done through using demand-side and supply-side rationing strategies (outputs). Rationing health services leads to rational utilization of health services (outcome). The impact of health services rationing depends on its effects on the equity of health and efficiency of health services. The information about reactions to health services rationing performance in each stage should be used for improvement.

The methodology of this study benefits from the comprehensiveness and theory generated nature. The product of the synthesis is not aggregations of data, but a theory grounded in the studies included in the review. The final proposed conceptual model breaks down the whole picture of rationing in the health system into discrete stages. The model can be applied to all health systems. However, this study had a limitation. The model placed too little emphasis on what is happening inside the health services rationing process at micro- levels, and it just provided a big picture of health care rationing at the macro- level of a health system.

Conclusion

Developing countries such as Iran face serious challenges in the limited resources and unlimited needs. These issues affect the effectiveness, efficiency, and equity of health care services provided to the society. This paper helps policymakers and managers understand the necessity of health services rationing and its mechanisms at various levels. Health services rationing could improve the utilization of health services and move toward universal health coverage. The stewardship of the health care system should regulate and balance rationing to control the behavior of the managers, health care providers (supply-side rationing), and patients (demand-side rationing).

Acknowledgements

This study was part of a PhD thesis supported by Tehran University of Medical Sciences (TUMS). The authors would like to thank the anonymous reviewers for their valuable comments on the earlier draft of the paper.

Conflict of Interests

The authors declare that they have no competing interests.

References

42. Dean M. London Rationing gets official seal of approval from UK Health Secretary. The Lancet. 2000;355(9197):49.