Chapter 2

Review and Synthesis of the Literature

2.1 Introduction

Findings from the review and synthesis of the literature are presented in this chapter. The purpose is to find out what is known about the geographic distribution of physicians, how geographic distribution is conceptualized and measured, and what problems have been encountered by researchers. This chapter lays the foundation for the detailed analysis of physician supply data in Chapter 3 by discussing the nature of the issues and by identifying the different approaches and methodologies used and the complexities in documenting physician distribution. In short, the literature review and synthesis guides and provides a rationale for the data analysis in Chapter 3.

Following this introductory section, there is a description of the methods used in the literature search, selection and review processes. Section 2.3 is an overview of the geographic distribution of physicians in Canada, the US and other countries based on a few published studies. Since a detailed analysis of physician distribution in Canada is found in Chapter 3, the purpose of this section is merely to show that the uneven geographic distribution of physicians is not a problem unique to Canada. Instead, many countries experience the same problem to a greater or lesser degree.

The remaining sections present the findings of the literature review. There are different ways to synthesize the literature. We have chosen to emphasize how the geographic distribution of physicians are understood by researchers and health services planners and how measurements of geographic distribution of physicians become increasingly more sophisticated. Sections 2.4 discusses the problems associated with counting the number of physicians and population and potential solutions to these problems. Section 2.5 presents a different approach involving geographic distances to describe the spatial dispersion of physicians and population. The last section of this chapter, Section 2.6, goes beyond describing where physicians are located. It discusses the feasibility of and difficulties in identifying physician "maldistribution" which involves comparing reality with service accessibility norms.

2.2 Methodology

As a review and synthesis of the research and health services planning literature on physician distribution is one of the two key components of this study, the literature search, selection and review processes are discussed in some detail as follows.

2.2.1 Literature Search and Review Process

Concerns over the uneven geographic distribution of physicians is widespread in Canada, an enormous country with a vast hinterland and a widely scattered population. It is, therefore, not unexpected that many studies have been produced on this topic. Similarly, there are many studies on physician distribution by researchers and health services planners in the US. Although the situations in the US are not identical to those in Canada, the concern over and responses to unequal access to medical care are similar in both countries. It is this sizeable body of literature from Canada and the US on the spatial distribution of physicians which is the focus of this review and synthesis.

One of the first major tasks was to identify, retrieve and select relevant studies for further review. The findings of those studies that were found useful were then analyzed and synthesized. In addition to the body of research noted above, there is a growing number of studies that are primarily of a policy analysis or program planning nature. These include studies that analyze or evaluate policies and programs that aim at addressing the uneven distribution of physicians. Many of these studies are done by ministries of health, health planning agencies, medical organizations and government task forces. Because most of these are not published in academic or professional journals and do not appear as monographs, the traditional approach of on-line keyword searches cannot be used to identify them. Furthermore, an increasing number of government and research documents now appear on the Internet, rather than in print format. Thus, we adopted a more comprehensive approach in gathering relevant studies. In addition to studies identified through on-line keywords searches, we included World Wide Web sites and unpublished reports by government ministries and other health care agencies. Also, CRaNHR researchers screened the references section of journal articles and books already in possession for other potentially useful studies. The entire process included the following major steps:

- Developed keyword search strategies and a list of keywords;
- Conducted on-line searches on such databases as Medline for potentially relevant studies using the identified keywords;
- Searched the Internet using various search engines;
- Canvassed selected experts, research centres, health planning agencies, and ministries of health for unpublished documents and other sources of material;
- Reviewed the references section of books, articles and reports for additional titles;
- Inputted titles of potentially useful studies and reports into Reference Manager;
- Screened titles and/or abstracts to identify studies for further review;
- Obtained hard copies of studies from the Laurentian University library, inter-library loan, the University of Western Ontario library (through the Northern Outreach Program library service) and various ministries of health, research centres and health care agencies;
- Reviewed studies to determine their relevance and usefulness;
- Created another database on Reference Manager to record selected studies; and
- Analyzed and incorporated useful studies in the synthesis of research findings.

The original intent, based on the Request for Proposals, was to select for review only those studies published in the 1990s. However, this time criterion would have severely restricted our ability to examine changes over time in the ways physician distribution were understood and measured. Thus, we decided to relax the time criterion in selecting studies. While some of the studies included in the literature review were published as far back as in the 1970s, the emphasis was on more recent publications.

More detailed information on the literature search and selection processes is contained in several appendices. Keywords used in on-line searches and the results of the searches are shown in Appendix I. Organizations and their uniform resource locators used in Internet searches can be found in Appendix II. Names of individuals and organizations successfully contacted are listed in Appendix III.

2.2.2 Scope of the Literature Review and Synthesis

There are different ways to classify this body of literature. Anderson and Rosenberg (1990), for instance, have proposed a 2-category typology: Descriptive and explanatory studies. Descriptive studies typically calculate and map population-to-physician ratios which are then compared and areas are classified as under- or over-serviced based on an ideal population-to-physician ratio. Explanatory studies take the number of physicians in an area or the population-to-physician ratio as the dependent variable and one or more factors are then used as independent variables in order to "explain" the distribution patterns. Joseph and Philips (1984) have further divided the explanatory studies into two categories: Those using an ecological approach and those using a behavioural approach. In addition, there is a large number of studies that describe, examine or evaluate various policies, programs and strategies designed to recruit and retain physicians to work in rural communities.

Because the literature on physician distribution is vast and due to the limited scope and time frame of the present project, it is important to specify at the outset what is and what is not included in this review and synthesis. Because Health Canada has commissioned a companion project that examines rural physician recruitment and retention issues, there is no need to duplicate the work of another team of researchers. Instead, the present project focuses on studies that *describe* the geographic distribution of physicians. Among studies that describe physician distribution, some focus on the problems of access to medical care in the inner core of major urban centres. Such studies are not included in the present review as the problems of access in inner cities, especially in the US, are not necessarily due to physician shortages (see, e.g., Eisenberg and Cantwell, 1976; Kaplan and Leinhardt, 1973; Pitblado, Pong and Jacono, 1995), but tend to be a problem related to poverty, ethnic minority status, cultural and/or linguistic differences and lifestyles.

The number of titles and abstracts reviewed and the number of publications selected for further review, as well as their sources, are shown in Table 2.1.

Table 2.1 Sources and Number of Publications Screened for Inclusion

Source	Number of Titles/Abstracts Reviewed	Number of Publications Selected for Further Review
On-line databases	2026	192
World Wide Web	41	26
Reference Section of Selected Books/Articles	214	145
Materials gathered from research centres, institutes and Government	32	12
Total	2313	379*

*Not all references selected were included in the final report

Although about 400 studies were selected for further review, not all of them were found to be useful and not all of those found useful were mentioned in the review and synthesis. Only those that fit the synthesis framework or those that have methodological or theoretical significance are mentioned in this report. The latter tend to be representative studies and are cited here for illustrative purposes.

While there are different ways to synthesize this body of literature, we have chosen to focus on the attempts by researchers to make the description of physician distribution more meaningful and a closer approximation of the reality. In other words, the purpose of this review and synthesis is to examine the pertinent literature and *to present the findings in terms of how the geographic distribution of physicians can be described more accurately and understood in more meaningful terms.* Unless the spatial distribution of physicians is described accurately, programs and policies that are based on it or directed toward it are unlikely to be effective. It is hoped that the results of our effort will be helpful to other researchers and health services planners who have to grapple with this seemingly simple but in fact very complex issue.

2.3 Physician Distribution: An Overview

Geographic distribution of physicians is just a shorthand way of describing what is of interest or concern to many people. Information solely about where physicians are located is of limited use. Besides, it is not particularly difficult to identify where physicians are located since this can easily be accomplished by plotting the postal codes of physicians' addresses on a map using geographic information systems techniques. Instead, people are mostly interested in *the spatial distribution of physicians relative to the spatial distribution of the population*. This is why the population-to-physician ratio is the most commonly used measure in research and in health services planning. It is a ratio that relates information on physician

distribution with information on population distribution. Besides being fairly easy to construct and understand, it is the basis upon which more elaborate and meaningful indices of spatial distribution, such as location quotients and Gini indices, are built.

What does the literature tell us about physician distribution in Canada, the US and other countries? In Canada, communities of up to 10,000 are often classified as rural. By this definition and according to the 1991 Census, 31.6% of the Canadian population lived in rural areas and 11.3% of physicians (18.6% of physicians in general or family practice and 3.8% of specialists) were located in these rural areas (Rourke, 1997). Major urban centres with 1,000,000 or more residents had 390 people for every physician. Outside urban centres, there were 1,175 people per physician.

The opinions of experts in Canada during the 1980s and early 1990s have been summed up by the authors of two studies. According to Horne (1986), most experts agreed that the country had an adequate supply of physicians, but the geographic distribution was skewed, involving both "maldistribution" between and within provinces and serious "shortages" in many rural and remote communities. Likewise, Barer and Stoddart (1992), in their landmark study on physician resources, have stated that "The consensus from our interviews was that there are serious problems with the geographic distribution of physicians in Canada. Indeed, this was one of the five 'first-tier' problem areas to emerge from those interviews" (p. 617).

Not surprisingly, the situations in many provinces mirror the national scene. For instance, commenting on physician resources in Nova Scotia, the Ministerial Task Force on Physician Policy Development (1993) has argued that "Per capita, Nova Scotia does not have a shortage of physicians, but rather one of the highest physician-population ratios in Canada. There is significant geographic maldistribution, however, both in family practice and in certain specialties" (p. 18). Similarly, according to the Ontario Council of Health (1983), the population-to-physician ratio in Ontario compared favourably with those in other provinces and other countries. However, inequalities appeared if the ratios were calculated on a regional or county basis.

Bergevin (1993) has stated that although Quebec had sufficient physicians in the 1980s, their geographic distribution posed a major problem. Thanks to various measures introduced by the provincial government to overcome this problem, there has been an improvement in the distribution of general practitioners and family physicians (GPs/FPs), but the situation for specialists has remained difficult in remote and intermediate regions. Earlier studies by Angus (1976) and Northcott (1980) in Alberta show that there was an uneven distribution of physicians in that province. Over the years, the growth in the number of physicians in urban centres outstripped that in rural communities. While there was a convergence trend with respect to GPs/FPs and primary care specialists (internists, obstetricians/gynecologists and pediatricians), the concentration of other specialists in major urban centres remained largely unchanged.

Despite differences in the two health care systems, the situation in the US with respect to physician distribution is not radically different than that in Canada. A study by the Office of Technology Assessment (1990) has reported that in the US, the number of active physicians per 100,000 population was more than twice as high in urban as in rural areas in 1988. Urban-rural differences were less pronounced for primary

care physicians than for other specialists. Within rural communities, population-to-physician ratios are related to size, with the smallest rural counties having less than one-half as many primary care physicians and about one-ninth as many nonprimary care physicians per capita as the largest rural counties. A rapid growth in physician supply notwithstanding and despite several studies by the Rand Corporation (e.g., Newhouse, et al., 1982; Schwartz et al. 1980; Williams, et al., 1983) which suggested that overall growth in the number of physicians would eventually solve physician shortage problems in smaller communities, rural-urban disparities have actually worsened during the last two decades.

In its third report, the US Council on Graduate Medical Education (1992) has come to a similar conclusion. It has identified access to medical care in rural areas as one of the major issues, despite a doubling in the number of physicians in the US in the 25 years prior to the publication of its report. Large metropolitan areas had three times the physician density of the smallest nonmetropolitan areas. The number of primary care shortage areas increased between 1980 and 1992 and an estimated 35 million Americans lived in these primary care shortage areas.

Moving from North America to the world stage, we find remarkably similar situations. Based on an analysis of health human resources carried out in many developed and developing nations, the World Health Organization Regional Office for the Americas (1987) has concluded that "(o)ne problem about which all the studies expressed concern is the poor geographic distribution of physicians, evaluated either by comparisons among various administrative regions or divisions in the countries, or between urban and rural areas" (p. 19). Other studies (e.g., Blumenthal, 1994; Mejia, 1987; Rosenthal and Frederick, 1984) have reported similar findings. Interestingly, even in countries where most or all physicians are employed in the public sector, geographic distributional problems are still common.

2.4 Beyond Head Counts

Simple population-to-physician ratios are used by many researchers and health services planners because they are simple to construct and easy to understand. As well, the data requirement is minimal compared to other measures of physician dispersion. A survey conducted by the National Ad Hoc Working Group on Physician Resource Planning (1995) of the Canadian Medical Association has found that nine of the ten provinces employed this ratio in their planning activities. Health services planners in Canada have used population-to-physician ratios extensively in medical workforce planning, beginning with the 1964 Royal Commission on Health Services (Roos et al., 1996a).

However, the pictures presented by population-to-physician ratios may be somewhat oversimplified or even misleading. This is because simple head counts of physicians and people within an arbitrarily defined area often obscure the underlying complexity of the situation. Many researchers have pointed out the inadequacies of the population-to-physician ratio including its failure to take into consideration differences in physician productivity, mobility of physicians and patients, physicians practising beyond the scopes of their specialties, substitution by other providers, the medical needs of the population, etc. (see, e.g., Angus

and Brothers, 1976; Connor et al., 1994; Contandriopoulos and Fournier, 1988; Hicks and Glenn, 1991; Roos et al., 1996a).

A more realistic assessment of the geographic distribution of physicians requires a better understanding of the conceptual and methodological aspects of measuring physician and population dispersion. The first step in advancing this area of research, as well as policy and program development in rural health care, is to enhance our ability to describe the geographic distribution of physicians more accurately and to take into account factors that impinge on care-providing and care-seeking behaviours.

The literature shows that researchers have adopted two different approaches in refining the measures of physician distribution. First, it is worthwhile pointing out that a typical population-to-physician ratio contains three pieces of information, both implicit and explicit: The *geographic area* within which the physicians and people are located, the *number of physicians* and the *number of people*.[1] Refinements of the ratio are attempts to conceptually clarify these three variables and to more accurately measure them. Second, some researchers prefer to avoid the need to define a geographic area altogether. Instead, the distributions of physicians and population are described in terms of distances between people and physicians. The rest of this lengthy section is devoted to discussing in greater details the first approach. The second approach, which uses distance measures, is discussed in Section 2.5.

2.4.1 Geographic Unit of Analysis

As Northcott (1980) has insightfully observed, one of the major issues in discussing equity in the distribution of health services or practitioners rests with the unit of analysis. Geographic unit of analysis refers to what geographic area is to be adopted in determining the numbers of people and physicians to be included in the population-to-physician ratio. Various units have been used by researchers. For instance, at the provincial level, both Anderson and Rosenberg (1990) and Coyte et al. (1997) use counties as the unit of analysis in their examinations of the distribution of physicians in Ontario. Northcott (1980) employs two geographic units of analysis in his study of physician availability in Alberta: Census divisions and incorporated cities,

---

[1] The population-to-physician ratio in an area is typically expressed as:

$$\frac{\text{Number of people residing in an area}}{\text{Number of physicians in an area}} = \text{number of persons per physician in an area}$$

The number of people residing in an area need not necessarily be the total population. For instance, if the denominator is the number of geriatricians, the numerator could be the number of people aged 65 and over in the area. Similarly, if the denominator is the number of obstetricians/gynecologists, the numerator could be the number of women in the area.

towns and villages. Pitblado and Pong (1995) examine variations among public health unit districts in access to physicians and dentists in Ontario. Contandriopoulos and Fournier (1988) use health planning regions to examine access to medical services in Quebec. It is also very common to compare rural with urban areas. Because of the importance of this unit of analysis, "rural" is discussed separately and in greater detail below. At the national level, health planning reports and health workforce data documents typically use provinces and territories as the unit for comparisons (e.g., Canadian Institute for Health Information, 1998b; Health Canada, 1992).

At the international level, Deliège (1987) and Mejia (1987), for instance, compare countries with respect to practitioner-to-population ratios. Ray (1987), on the other hand, compares groups of countries at different levels of economic development with respect to the numbers of physicians and nurses per 100,000 population.

The major problem of using administrative areas, such as counties, or statistical enumeration areas, such as census divisions, is that they are "artificial" and not necessarily the most appropriate unit of analysis. Both Jacoby (1991) and Kindig and Ricketts (1991) have urged the adoption of units of analysis that are more appropriate for the analysis of physician services, such as "physician market area" or "health service trade area".

Researchers have experimented with more appropriate geographic units of analysis. Makuc et al. (1991) have attempted to define national "health service areas" in the US. A health service area is an area with one or more counties that are relatively self-contained with respect to the provision of routine hospital care. Using cluster analysis, the authors are able to group all counties into 802 health service areas. These areas are more self-contained with respect to the provision of hospital services, thereby providing a more appropriate geographic unit than the county for measuring the availability of health care. The median number of patient-care physicians per 100,000 population is 84 for nonmetropolitan health service areas and 132 for metropolitan health service areas.

Krasner et al. (1977) has advocated using Zipcode Sectional Areas as a geographic unit for analyzing the distribution of physicians in the US. These areas are identified by the first three digits of the 5-digit Zip code number. Because Zipcode Areas were established on the basis of local transportation patterns, they tend to approximate trading areas. Krasner et al. have examined the distribution of dermatologists in the US based on this unit of analysis. Using a similar rationale, Morrow (1977) has promoted the use of Office of Business Economic (OBE) areas as the unit of analysis. OBE area designations are based on factors describing commuting patterns exhibited by people travelling to receive or sell their services and products in the US. According to Morrow, OBE areas are better approximations of health care market areas. In Canada, Roos et al. (1996b) have used "physician service areas" to study the needs for generalist physicians in Manitoba.

"Rural" as a Geographic Unit of Analysis

Many studies compare rural with urban areas with respect to physician availability. Joseph and Philips (1984) are correct in pointing out that macro-scale studies of physician availability have focused on rurality as an important attribute for the identification of geographic disparities in medical human resources. In such studies, the unit of analysis is rural versus urban or rural versus non-rural. Instead of comparing physician availability in many units, researchers typically present physician distribution data for rural and urban regions with a view to showing the differences between them. In addition, some researchers have introduced the concepts of "frontier" and "remote" in order to distinguish the sparsely populated rural areas from those that have more inhabitants.[2]

However, there are almost as many definitions of "rural" as there are researchers. To make a bad situation worse, it is not uncommon for authors to use the term rural without specifying what it refers to. Bosak and Perlman (1982) have reviewed 178 rural mental health and sociology articles and found that 43% of them do not include a formal definition of rural. Similarly, Ricketts and Johnson-Webb (1997) have reviewed articles on physician practice locations and physician recruitment and retention issues published in the *Journal of Rural Health* between 1993 and 1995 and have found wide variations in how rural was defined. Johnson-Webb et al. (1997) have noted that in the US, policy-makers in federal, state and local governments have not reached a consensus on what rural means. Inconsistent definitions often exist within an agency, such as the Health Care Financial Administration, which uses different designations of rural for different legislation-mandated programs. Furthermore, the concepts of rural and urban are not constant. Bollman and Biggs (1992), for instance, have documented the changing definitions of rural and urban as used by the Canadian Censuses of Population over time.

In its report, the Advisory Panel on the Provision of Medical Services in Underserviced Regions (1992) of the Canadian Medical Association defines rural communities as those with a population of 10,000 or less. This definition appears to be similar to the definition of "rural and small town Canada" used by Statistics Canada (Mendelson and Bollman, 1998). "Rural and small town Canada" refers to the population living outside the commuting zones of larger urban centres - especially outside Census Metropolitan Areas (with population of 100,000 or more) and Census Agglomerations (with core population of 10,000 - 99,999). Statistics Canada (1997; see also Table 2.2 and Chapter 3) classifies areas into five categories: urban core (Census Metropolitan Area/Census Agglomeration), urban fringe (urban areas within CMA/CA boundaries but not contiguous with the urban core), rural fringe of CMAs/CAs, urban outside CMAs/CAs, and rural. In their study, Sanmartin and Snidal (1993) define rural physicians as doctors living in areas with

---

[2]The Rural Committee of the Canadian Association of Emergency Physicians defines "rural remote" as rural communities that are 80-400 km, or about one to four hours of travel in good weather, from a major regional hospital; and "rural isolated" as rural communities greater than 400 km from a major regional hospital (Rourke, 1997). In the US, the National Rural Health Association defines frontier counties as counties with less than six persons per square mile (Weinert and Boik, 1995).

a “0" as the second digit in the postal code. An agreement between the Ontario Ministry of Health and the Ontario Medical Association defines communities that have fewer than 10,000 people and are at least 80 km from an urban centre with 50,000 population or more as “specified” or “isolated” communities (Rourke, 1997). This is often used as a definition of rural in health services planning in Ontario. Leduc (1997) has proposed the General Practice Rurality Index.

Each of these definitions has its attractiveness and limitations. As an illustration, Wilkins (1993) has discussed the potential of using the postal codes of the addresses of patients and health care providers in the spatial analysis of health services provision. On the other hand, the shortcomings of the Canada Post definition of rural (i.e., areas with a “0" as the second digit in the 6-digit postal code) have been noted by Wootton (1996). Some of the major Canadian definitions of “rural” are summarized in Table 2.2.

In the US, the two most common designations used in health care delivery for classifying rural populations are the county-based “metropolitan” and “non-metropolitan” designations employed by the Office of Management and Budget and the “urban” and “rural” designations used by the Bureau of the Census. “Non-metropolitan” counties are those counties without a city of 50,000 or more residents, or counties without an urbanized area of 50,000 or more and without a total population of 100,000 or more. “Rural” populations are those living outside of urbanized areas in towns of fewer than 2,500 persons or in open country (US Department of Health and Human Services, 1992). The problem with these definitions is that a sizable percentage of people in census-defined rural areas live in Office of Management and Budget-defined metropolitan areas, and, conversely, a considerable percentage of metropolitan residents live in census-defined rural areas (Goldsmith et al., 1998). In order to overcome the problems of using the conventional definitions of rural, Weinert and Boik (1995) have designed the Montana State University Rural Index. This index assigns a value (degree of rurality) to each household on the urban/rural continuum using only two variables, population of the county of residence and distance to emergency care. In addition, Fickenscher and Lagerwey-Voorman (1992) have differentiated four types of rural areas: Adjacent rural areas, urbanized rural areas, countryside rural areas and frontier areas.

Leduc is right when he opines that “(a) widely accepted and validated definition of ‘rural’ in the context of medical practice has not yet been developed in Canada” (p. 125). Perhaps, as Halfacree (1993) has suggested, the search for a single, all-purpose definition of rural is neither desirable nor feasible. How rural should be defined depends on the task at hand. The problems facing researchers and health services planners is the inability to compare studies and findings with respect to physician availability in rural areas since the term is not used in a consistent manner. Thus, rural as a geographic unit of analysis is quite restricted in its utility. As it appears in most studies, the concept is either implicit or not consistently defined. Also, because “rural” covers a vast territory and is not sufficiently differentiated to reflect varying degrees of rurality, it tends to mask intra-regional variations in physician availability and population dispersion. However, a rural-urban comparison can be effective in highlighting or underscoring gross unevenness in physician distribution within a jurisdiction. Its imprecision notwithstanding, the term “rural” is unlikely to disappear from everyday parlance or from policy and planning discourse since it is such a convenient label and is so ingrained in our consciousness.

Table 2.2 Selected Methodologies Used to Define "Rural" in Canada.[1]

	Principal Defining Criteria		
Organization (Source)	**Population Characteristics**	**Distance Characteristics**	**Other/Comments**
Statistics Canada (Statistics Canada, 1997; and previous Census Dictionaries published by Statistics Canada)	<1,000 population <400 persons/sq.km.	adjacency to Census Metropolitan Areas (CMA) and Census Agglomeration Areas (CA) used to provide sub-categories	Accordingly, each enumeration area of Canada classified into: C urban core C urban fringe C rural fringe C urban outside CMA/CA C rural outside CMA/CA
Canadian Medical Association and Canada Post (see: Wilkins, 1993; Sanmartin and Snidal, 1993; or Wootton, 1996)			Second digit of postal code "0" (see discussion in Chapter 3)
Research Sub-Committee of the Interdepartmental Committee on Rural and Remote Canada (1995) using the criteria of the Organization for Economic Co-operation and Development. (See also: Canadian Rural Information Service of Agriculture Canada, 1999)	<150 persons/sq.km.	adjacency to metropolitan areas northern hinterlands used to provide sub-categories	Accordingly, each census consolidated subdivision (CCS) classified into: C agglomerated C intermediate C Rural, metro adjacent C Rural, non-metro adjacent C Rural north
Ontario Medical Association (Rourke, 1997; OMA, www.oma.org)	<10,000 population	distances to a community of 50,000+ used to provide sub-categories	Used to define groups of physicians re. continuing medical education subsidies: Group 1 > 80 km Group 2 50-80 km

Table 2.2 (continued)

	Principal Defining Criteria		
Organization (Source)	**Population Characteristics**	**Distance Characteristics**	**Other/Comments**
Canadian Association of Emergency Physicians (1997)	<10,000 population	distances from a major regional hospital used to provide sub-categories: <80 km or 60 minutes 80-400 km or 1-4 hours >400 km or >4 hours	 Rural close Rural remote Rural isolated
Society of Rural Physicians of Canada (Leduc, 1997)	drawing population (i.e. the total number of people) of communities	• distance from closest advanced referral centre • distance from closest basic referral centre • number of GPs within 25 km of community • number of specialists with 25 km of community • presence of an acute care hospital with 25 km of community	The population and distance criteria are weighted to generate a score that is labeled as the General Practice Rurality Index.
Rural & Small Town Programme, Mount Allison University; prepared for the Canada Mortgage and Housing Corporation (Ashton et al., 1994)	C Omitting CMA/CA communities as defined by Statistics Canada, divide Census Subdivisions into incorporated and unincorporated categories C compute means and standard deviations of three variables: total population, population density, % workforce in primary industry	adjacency to CMA/CA used to provide sub-categories of unincorporated communities	The number of standard deviations away from the mean of the three population criteria and the adjacency index added together to produce a score for each non-CMA/CA community to generate an 8-fold classification: Rural (four categories) Small Town (four categories)

[1] The methods included here are not wholly designed for use in defining rural for the purposes of health research or policy formulation.

Patient and Physician Mobility

As Krasner et al. (1977) and Morrow (1977) have alluded to, the choice of geographic units of analysis is complicated by another factor, namely, patient and physician mobility (Hong and Kindig, 1992; Makuc et al., 1991; Wing and Reynolds, 1988). As peripatetic creatures, patients and physicians are not necessarily constrained by administrative or geopolitical boundaries in seeking or delivering care. Trips to another community to seek or to deliver care are commonplace. Kleinman and Makuc (1983) found that the percentage of physician visits that involved county-border crossing in the US varied from 7% for people in greater metropolitan core counties to 45% for those residing in rural counties. Similarly, Contandriopoulos and Fournier (1988) noted the substantial interregional mobility of physicians and patients in Quebec. In their study of physician visits by older persons in a retirement community in northern Ontario, Pong, Salmoni and Heard (in press) found that while over 95% of the visits to GPs/FPs took place within the community, over two-thirds of the specialist visits took place in other cities, which required patients to travel long distances.

It is worth noting that several provinces such as British Columbia, Manitoba and Ontario have travel subsidy programs, such as the Northern Health Travel Grant Program in Ontario, to assist patients in rural or northern communities who have to travel long distances to seek medical care. The federal government has a similar program for status Indians on reserve (Northern and Rural Health Task Force, 1995). As well, in some provinces such as Ontario, there are programs to bring specialists to remote or rural communities to provide specialist care that is not locally available. The Saskatchewan Physician Resource Planning Task Force (1994) has voiced it support for visiting specialist services. Such programs make border crossing for the purpose of seeking or delivering medical care more affordable and more likely.

What are the implications of border crossing by patients and physicians? As mentioned earlier, the conventional population-to-physician ratio is based on the assumption that medical care provision and consumption take place within a defined geographic area. This assumption is untenable due to patient and physician mobility. As Wing and Reynolds (1988) have correctly observed, "medical service areas in *any* health care system allowing freedom of choice for patients are not discrete, self-contained units, they are probabilistic and interpenetrating" (p. 652; original emphasis). The simple population-to-physician ratio may provide inaccurate information because the physicians represented in the denominator of the ratio for a particular geographic area typically serve more people than those counted in the numerator. On the other hand, the people counted in the numerator may receive care from physicians outside the area where they reside.

Researchers inevitably find themselves in a Catch-22 situation when deciding what geographic unit of analysis to adopt. If the chosen unit of analysis is small, the measures of physician availability are likely to be distorted by patients' and physicians' cross-border travels. Contrariwise, if the unit of analysis is very large like country, province or "rural", while people are less likely to journey beyond borders, the measures of physician distribution tend to hide intra-regional differences (Anderson and Rosenberg, 1990; Hadley, 1979). Because of this analytical dilemma, Connor, Kralewski and Hillson (1994), in their review of the

literature, have differentiated two types of analysis. One type is called the "contained-area model" which is based on the traditional concept of a self-contained local service area with well-defined boundaries, within which a local practitioner or group of practitioners provides most of the care for local residents. The other is the "full-travel model" which measures the effects of travels on physician access as a continuous function of distance or time. It allows each of several practitioners to have a portion of the market for a given population.

A number of analytical approaches have been suggested with a view to overcoming the difficulties posed by geographic unit of analysis and/or travels by patients and physicians. For example, the Quebec Ministry of Health and Social Services (undated) has proposed a physician resources distribution plan which estimates physician availability based on full-time equivalents (FTEs).[3] The objective of the plan is to measure regional disparities in terms of availability and access to medical care. Data can be examined in several ways:

- Number of FTE physicians *established* in a region: This refers to the actual production of services by physicians within a region, regardless of the origin of the patients receiving these services or the region in which they are provided.
- Number of FTE physicians *available* in a region: This refers not to the region where the physicians are established, but to the region in which services are provided.
- *Consumption* in FTE physicians by the population of a region: This allows the examination of the volume of medical services consumed by the population of a region, translated into FTE physicians, regardless of the regions in which the services are consumed or the origins of the physicians providing them.

The difference between the number of FTEs established and the number of FTEs available provides an indication of the *extent of itinerancy*. On the other hand, the difference between the number of FTEs established or available in a region and the population's consumption, in terms of FTEs, provides an indication of the region's *self-sufficiency* in the provision of medical care.

Wing and Reynolds (1988) have suggested another approach that does not rely on non-overlapping geopolitical units. They have developed implicit physician service areas that are probabilistic and interpenetrating in order to yield estimates of the supply of physician services in small geographic areas. This is accomplished by allocating a portion of the services of each physician to his/her home area and to other areas in proportion to both the availability of potential patients and the propensity of patients to travel for medical care. The final estimate of the availability of physician services in each small area is the sum of the service proportions of every physician in all of the small areas. The total supply of physician services is the same as the original total, but the distribution is adjusted to reflect the time that patients are willing to spend travelling to obtain medical care.

---

[3] A more detailed discussion of full-time equivalents can be found in Section 2.4.2 and Chapter 3.

2.4.2 Counting Physicians and Patients

Another important step in the process to improve the population-to-physician ratio is to ensure a more accurate enumeration of medical practitioners. This is more than just counting all physicians in an area, which is usually not a very difficult task since in Canada all physicians have to be registered with a provincial or territorial college of physicians and surgeons in order to practise medicine. By checking the list of registered physicians, one can get a fairly accurate count of the number of physicians in an area. The much more arduous task is to find out what they do and how much they do. In other words, it is necessary to determine if the physicians are more or less identical in the production of clinical services. The simple population-to-physician ratio does not usually take into consideration non-clinical work, productivity, varying activity levels and the influence of age or sex on workload and practice pattern. As a result, the ratio may over- or under-estimate the supply of clinical services by physicians. In the final analysis, it is not the number of "warm bodies" that matters. What is important is the production of medical care.

A number of studies have documented the gradual decline in the number of hours Canadian physicians work. For example, according to the 1990 survey on physician resources conducted by the Canadian Medical Association, physicians reported working on average 4.1 fewer hours per week in total activities in 1990 than in 1986, and 5.7 fewer hours per week than in 1982. For GPs/FPs, the figures are 44.8 hours in 1990, 46.9 hours in 1986 and 51.5 hours in 1982. This represents a 13% decrease in hours of work in less than 10 years. Specialists reported working 48.5 hours per week in 1990, 4.6 hours less than in 1982 (Sanmartin and Snidal, 1993). For all intents and purposes, less hours of work, unless counterbalanced by enhanced efficiency, means less human resources available. Surprisingly, of the studies reviewed which examine changes in population-to-physician ratios over time, none has factored the decline in practice hours into the calculations.

In Canada, as well as in many other industrialized nations, the proportion of women physicians in the medical workforce is rising steadily. Currently, about half of the students in Canadian medical schools are women. Many studies (e.g., Contandriopoulos and Fournier, 1983; Ferrier et al., 1988; St-Laurent-Gagnon et al., 1993; Woodward and Hurley, 1995) have shown that female physicians tend to work less hours per week and less weeks per year than their male counterparts. The implications for physician resources planning and medical practice of the changing sex composition in the Canadian physician workforce have been examined by Reamy and Pong (in press) and Williams et al. (1993). It is also known that as physicians approach retirement age, they tend to reduce their workload. Additionally, physicians differ in their involvement in non-clinical work such as research, teaching and administration. For example, it has been well documented that physicians affiliated with medical schools tend to spend more time on such activities. Results of the 1997 National Family Physician Survey commissioned by the College of Family Physicians of Canada show that 83.2% of a GP/FP's time was spent in direct patient-care, while 16.8% was spent on related activities including research, teaching and administration (Irvine and Pong, 1998). Therefore, when examining physician availability, it is necessary to consider the effects of age, sex and type of activity on the production of clinical services.

Another issue that needs to be considered is specialty mix. Werner, Langwell and Budde (1979) have raised concerns about possible errors associated with the use of a single population-to-physician ratio that includes different specialties. The question is whether the physicians included in the denominator of the ratio are sufficiently homogeneous in terms of specialty to permit addition without introducing distortions. If the specialties included are heterogenous and if different regions have different specialty combinations, the ratios may not be comparable. Their research shows that primary care specialties (i.e., general and family practice, internal medicine, pediatrics, obstetrics/gynecology and general surgery) differ significantly in terms of total hours of work, allocation of time to different activities, productivity and number of patient visits. The authors have suggested making necessary adjustments to the numerator, like using weighted averages, as a refinement of the population-to-physician ratio.

A related issue is specialty substitution which refers to physicians in one specialty providing clinical services that are typically rendered by physicians in other specialties. This is possible because the scopes of practice of different specialties overlap to a considerable extent. Specialty substitution is particularly prevalent in rural communities where specialists are few and far between. Rural GPs/FPs may perform some medical procedures that, in major urban centres, would typically be done by specialists. Similarly, general surgeons may perform some orthopedic surgeries if orthopedic surgeons are not available locally; internists may do some cardiology work, if there are no cardiologists; etc. (see, e.g., Roos et al., 1996a, regarding substitution in general surgery). The Saskatchewan Working Group on Physician Need (1997), for example, has acknowledged that non-certified specialists have been a main source of supply of specialist services in that province, particularly in communities other than Regina and Saskatoon. Very few of the studies reviewed have taken the issues of specialty mix and specialty substitution into account when measuring physician distribution.

The Full-time Equivalents Approach

The above discussion shows the complexity in the seemingly simple task of counting physicians. One approach that has been introduced in order to more accurately reflect the availability of physician resources in an area is the use of the full-time equivalents (FTEs) technique. The use of FTEs is increasingly common in research studies and planning documents, suggesting that more and more people realize the shortcomings of head counts and the need to use more meaningful measures (e.g., Expert Panel on Physician Resources, 1996; National Ad Hoc Working Group on Physician Resource Planning, 1995; Roos et al., 1996b).

The national FTE methodology was used by the Federal/Provincial Working Group for the Development and Review of Medical Care Statistical Indicators in 1989 and was seen to be a more appropriate measure of physician activity. The original calculation of national FTEs were based on fee payment data which

captured the number of services provided and the fee schedules.[4] Because the calculation relied on payment data, the FTEs were problematic in a number of ways. For instance, radiology and laboratory specialists were excluded because most of them were not on fee-for-service. Similarly, other non-fee-for-service physicians and some clinical services that were not funded through fee-for-service were not included in the FTE estimates.[5]

New Brunswick has introduced a modified FTE methodology with a view to overcoming some of the problems in the national FTE methodology. The New Brunswick methodology is more comprehensive in nature as it includes radiology and pathology and non-fee-for-service payments which represent about 11% of physician payments in that province. However, some limitations remain. For example, specialty designation is based on certification not actual profile of practice; uncertified specialists are not captured and all non-clinical services such as community medicine, hospital administration, research are not captured (New Brunswick Department of Health and Community Services, 1992; Reamy and Pong, in press).

The Advisory Panel on the Provision of Medical Services in Underserviced Regions (1992) has used both head counts and FTEs in its analysis. It has found that in 1989/90, 16.8% of GPs/FPs were in rural communities. But this percentage was 18.2% when FTEs were used. Since the rural physician FTE count is higher than the number of GPs/FPs, this suggests that physicians in rural practice are providing a greater number of services than average or that they were working longer hours than average. The New Brunswick Department of Health and Community Services has also used the FTE methodology in its physician workforce planning. A comparison of FTEs to the traditional head counts reveals that the 1,078 FTEs in 1989/90 represented 1,123 physicians, a difference of 4%.

Determining Population Size

Even more difficult than counting physicians is counting patients or potential patients. Again, the problem is not so much the enumeration of people in a defined geographic area as determining the medical care requirements or utilization of the population. It is well known that certain population groups, such as women and older persons, consume more health services than others. It is important, therefore, to make sure that the numerator in the population-to-physician ratio adequately reflects the service consumption patterns of

---

[4] A more detailed discussion of the national FTE methodology can be found in the report of the National Ad Hoc Working Group on Physician Resource Planning (1995). Also, Appendix VI outlines the FTE calculation as published by the Canadian Institute for Health Information (1998a).

[5] Major improvements in the FTE methodology are expected as the Canadian Institute for Health Information will soon be collecting data on non-fee-for-service physicians for those provinces able to provide such data. This would enable, for instance, the inclusion of radiology and laboratory specialists in the FTE calculation. In addition, it will change the base year from 1985/86 to 1992/93 and revise the benchmarks used in the FTE methodology in order to measure changes in fees overtime.

the population, not just the population size. Two communities with the same number of residents but with different demographic structures and health characteristics could have very different medical care needs or utilization patterns. This is particularly pertinent when analyzing rural and northern regions where the health status of the residents is generally below the national or provincial norm (see, e.g., Braden and Beauregard, 1994; Fickenscher and Lagerwey-Voorman, 1992; Kohrs and Mainous, 1995; Northern and Rural Health Task Force, 1995; O'Neil, 1989; US Department of Health and Human Services, 1992; Weller, 1981; Young, 1988).

Realizing this, the National Ad Hoc Working Group on Physician Resource Planning (1995), among others, has recommended several improvements to the population-to-physician ratio methodology, including adjusting the population for age-sex composition. Age-sex standardization can be seen as an initial step in adjusting the numerator of the ratio. Additional adjustments are possible, such as by taking into account variations in health status or morbidity among populations, even though the additional data requirements could be substantial.

The Expert Panel on Physician Resources (1996), in its report on physician workforce planning in Ontario, has adjusted the population component of the population-to-physician ratio by factoring into the calculations all-cause standardized mortality ratios as a proxy for morbidity.[6] Roos et al. (1996a) used three approaches, including the population-to-general surgeon ratio approach and the population needs-based approach, to analyze the needs for general surgeons in southern Manitoba. The needs-based approach took into account the age structure and various health characteristics (e.g., premature death rate, indices of general health and physical function and a socioeconomic risk index) of the population. Using the ratio approach, the researchers found that more general surgeons were needed. However, the needs-based approach did not support the necessity to recruit more general surgeons.

All this represents a series of attempts to refine and elaborate the population-to-physician ratio for research and planning purposes. Ideally, efforts should be made to modify both the numerator and the denominator of the ratio in order to more accurately reflect physician distribution, as well as medical care provision and consumption, as exemplified by a study by Coyte et al. (1997). In this study of the availability of GPs/FPs in Ontario, Coyte and his associates have introduced the notion of "physician density" which is the ratio of practice-intensity equivalent GPs/FPs to use-intensity equivalent residents. A "practice intensity equivalent index" is calculated, using age-specific OHIP fee service claims in 1990. Not unlike the FTE methodology discussed earlier, this index affords an opportunity to adjust the supply of GPs/FPs to reflect variations in

---

[6] The need to take the health status and/or health care needs of a population into consideration is also implicit in the definition of the Index of Medical Underservice (IMU). The IMU is used by the US federal government to define Medically Underserved Areas. Unlike other measures of underservicing which tend to rely solely on the number of physicians and population size, the IMU adds to the primary care physician-to-population ratio a measure of poverty in the area, the age structure of the population and a crude measure of health status or outcome (Office of Technology Assessment, 1990; Taylor et al., 1994).

service provision. Similar methods are used to adjust the size of the population for patterns of health care use. While there was on average 1 GP/FP per 1,000 people in 1990, densities for individual counties ranged from a low of 0.33 in the Sudbury District to a high of 1.74 in Frontenac. After adjusting for both practice intensity and the population's utilization patterns, the revised densities ranged from a low of 0.35 in the Sudbury District to 1.61 in Frontenac. Although considerable disparities in GP/FP availability persist, the difference has narrowed by more than 10% after adjustments.

2.5 Distance Between Physicians and Population

We have mentioned the problems associated with using the population-to-physician ratio. The two most critical problems are the need to adopt an often artificially defined geographic unit of analysis and the assumption that there is no mobility across regional boundaries on the part of service providers and service users. These problems have presented many conceptual and analytical difficulties. For this and other reasons, some researchers interested in the geographic distribution of physicians, as well as other types of health services, have adopted a different approach by using distance between physician and patient or potential patient as the analytical focus.

When access to medical care is the main concern, a more direct measure, and probably a measure that is more meaningful to most people, is the distance between service consumer and service provider or the amount of time needed to travel in order to access service. The issue is no longer the number physicians in an area or the number of communities without a physician, but how far an individual has to travel or how much time he/she has to spend on travelling in order to access service. Thus, the level of analysis has shifted from the geographic area to the individual (Jacoby, 1991; Shannon et al., 1969; Williams et al., 1983). Nonetheless, as explained by Joseph and Bantock (1984), these two types of measure - population-to-physician ratio and distance between population and physician - are complementary. While the former are indicators of broad supply and demand relationships across large regional units, the latter can be used to gauge the variability in potential access within such large regions.

The how-many-miles-to-the-doctor approach is not entirely new. Using data from the 1971 census, Angus and his associate (Angus, 1976; Angus and Brothers, 1976) examined the geographic distribution of physicians in Alberta by using the "proximity of population to physicians" method which was an attempt to establish a correlation, in terms of distance, between people and the physician closest to them, irrespective of the boundaries of regions or counties within the province. The researchers examined the average distance that populations, using the Enumeration Area (EA) as a base, would have to travel in order to obtain medical care. The distance was calculated between the centre of the EA and the centre of the municipality in which the physician was located. This was accomplished by adding a special Universal Transverse Mercator Coordinate System to the 1971 Census of Population summary tape file. They found that in 1971, for the most part, Albertans should be able to access physician services in their respective Census Divisions.

Williams et al. (1983) analyzed the distances that rural residents in 16 states in the US had to travel to

receive medical care from various types of specialists and documented changes in physician availability in the 1970s. In order to estimate the distance to the doctor, the authors calculated the distance (as the crow flies) between a point corresponding to the centre of each 33-square-mile grid and a point corresponding to the latitude and longitude of the nearest physician of a designated specialty. In 1970, only 13 rural residents in 100 lived more than 10 miles, and only 2 in 100 lived more than 20 miles, from a practising physician in the 16 states examined. Actual driving distances averaged about 20-25% longer than the straight-line distances. Between 1970 and 1979, there were across-the-board reductions in distance to a physician for most specialty categories. By 1979, approximately four-fifths of rural residents were within 20 miles of specialists in internal medicine, general surgery, obstetrics/gynecology and pediatrics, and fewer than 5% were more than 50 miles from such specialists.

Joseph and Bantock (1984) used a distance indicator to study the changes in potential accessibility to GPs/FPs in Bruce and Grey Counties from 1901 to 1981. They found a decline and a centralization of GP/FP services in this southern Ontario rural region during the 80-year period. Individual rural localities fared differently, dependent mostly on their location relative to the central place network. The authors concluded that the concentration of GPs/FPs in larger settlements had put the dispersed rural population at a distinct disadvantage in terms of access to medical care. In their study of accessibility of medical services in the Abitibi-Temiscamingue region of Quebec, Thouez et al. (1988) also used a measure of geographic access which incorporated the notion of distance between physicians and population. They found that between 1973 and 1982, there was an improvement in the potential accessibility of GP/FP services in the region, due to a marked increase in the number of GPs/FPs, rather than a reallocation of physician supply. With respect to specialists, the improvement in potential accessibility was less pronounced, despite a sizeable increase in the number of specialists, because there were less settlements with specialists. Thus, both the Joseph and Bantock study and the Thouez et al. study have found a growing centralization of physicians in larger settlements in rural areas over time.

The most recent attempt to examine the proximity of the population to physicians in Canada is a study by Ng et al. (1997). Proximity between population and physician is based on the aerial distance to the nearest physician from a representative point within each of Canada's 45,995 enumeration areas. Since the majority of Canadians live in urban areas, close to 87% of Canadians were less than 5 km from a physician in 1993. Nearly all (99%) residents of large urban centres (with 1,000,000 population or more) were less than 5 km from the nearest physician. But in non-urban areas, only 56% of residents were located that close to a physician. Not surprisingly, the smaller the community, the farther the distance to the nearest physician. In British Columbia, Ontario and Quebec, 91% of the residents were less than 5 km from a physician. By contrast, in Nova Scotia, despite the high population-to-physician ratio, only 70% of the residents were less than 5 km from the nearest physicians. And in Saskatchewan, Prince Edward Island and New Brunswick, less than 70% were no more than 5 km away. In 1993, 86% of Canadians were within 5 km of a GP/FP, but distances to specialists varied widely, depending on the specialty. Generally speaking, the smaller the specialty, the greater was the distance to the nearest specialist.

2.5.1 Distance-to-Physician Approach as Planning Tool

The proximity-of-population-to-physician approach can be used in health service planning. Because the traditional population-to-physician ratio relates people to doctors in a defined geographic area, resulting in an indicator highly dependent on artificial boundaries, Angus and Brothers (1976), as well as others, have advocated the use of the proximity approach which is not constrained by geographic units such as county, district or census division.

GMENAC has recommended that five basic types of health services should be available within some minimum time standards: 30 minutes of travel for emergency medical care, 30 minutes for adult medical care, 30 minutes for child medical care, 45 minutes for obstetrical care and 90 minutes for surgical care. The 30-minute travel time for non-emergency medical care has become the standard in health services planning in the US (Jacoby, 1991). In Canada, the Advisory Panel on the Provision of Medical Services in Underserviced Regions (1992) has recommended that "for the purpose of medical resource planning, an 'underserviced region' be defined as a geographic area where access to local health services is justified but where such services have not or cannot be maintained consistently. The guideline for temporal access referred to above should be used to identify such areas" (p. 35). In terms of temporal access to the various levels of care, the Panel has suggested that primary care should be available within 30 minutes, secondary care within two hours and tertiary care within five hours.

Saskatchewan Health has developed a physician resource planning approach which assumes that residents will receive a particular service in the closest community where such service is available. This approach is referred to as the "closest-to" population analysis. For any particular "closest-to" population analysis, it is necessary to first identify those communities which will provide the particular service. The next step is to assign the population to the provider community which is closest to that population in terms of travel time. This results in the calculation of a "closest-to" catchment area population for each provider community. The specialist supply is then distributed to each provider community in proportion to that community's "closest-to" population. If only three centres (Saskatoon, Regina and Prince Albert) are designated to provide specialist services, 28.7% of the population in 30 districts would have to travel 90 minutes to the closest one of these three centres. And 19.7% would have to travel more than two hours to the nearest centre. If the number of provider centres is increased to six (Saskatoon, Regina, Prince Albert, North Battleford, Yorkton and Moose Jaw), 16.3% of the population would have to travel more than 90 minutes to the closest of these centres. Only 7.5% would have to travel more than two hours to the nearest centre (Saskatchewan Working Group on Physician Need, 1997).

2.6 Identifying Physician Maldistribution

Thus far, we have studiously avoided using the term "*mal*distribution" when describing the geographic distribution of physicians (except when quoting other authors). Instead, we have chosen to use the more prosaic phrase "uneven distribution of physicians". This is because, as Ginzberg (1978) has cogently

pointed out, unless there is a concept of acceptable distribution, the prefix "mal" does not make sense. Something is bad only in relation to something that is not bad. However, there is strong temptation to move from "uneven distribution" to "maldistribution" in analyzing physician distribution. Not content with merely describing the geographic distribution of physicians, many studies in the body of literature being reviewed make comparisons with respect to the availability of physicians, with a view to identifying doctor-shortage areas or maldistribution. There are two main types of comparison: Comparing the population-to-physician ratio in one area with an ideal or optimal ratio and comparing the ratios in two or more areas.

2.6.1 Comparing Observed Distribution with "Optimal" Distribution

For both research and planning purposes, it is often useful to be able to identify areas or communities as underserviced or experiencing a maldistribution problem. The most common method used is to compare the observed or actual distribution of physicians with some standard which can be an "ideal" population-to-physician ratio, an "optimal" ratio recommended by some experts or authoritative bodies, a "required" number of physicians or an "objective" standard of adequate medical care.

If the number of physicians required to serve the population of an area can be ascertained, it could enhance health service planning capability. The required number of physicians, along with other data, can then be used to construct various indices. As an illustration, the ratio of the number of FTE physicians who reside in an area to the number of physicians required constitutes an *index of distribution* of medical resources. The ratio of the number of FTE physicians serving a population to the number required is an *index of access to services*. Using these indices, Piché (1993) has examined physician resources in various regions of Quebec. With a distribution index of 1.56, the Montreal region has 56% more FTE specialists than are required to serve its population. The Lanaudière-Laurentides and North Shore regions, by contrast, have the fewest FTE specialists with only 40% and 43%, respectively, of the numbers required. On the other hand, regional disparities are considerably less pronounced with respect to GPs/FPs.

In the US, various measures of underservice have been introduced by the federal government. For example, the US Bureau of Health Manpower has defined the Critical Medical Shortage Area as an area that has a ratio of resident population to FTE non-federal, primary care physicians greater than 4,000:1 (Hadley, 1979). The Health Professional Shortage Area designation is used to identify locations in which to place physicians under the National Health Services Corps Program. A county, a sub-county unit or an aggregation of counties is seen as underserviced if the population-to-primary care physician ratio in the area is greater than 3.500:1 (Taylor et al., 1994).

The Medically Underserved Area (MUA) designation is used by the US federal government to identify areas where health maintenance organizations could receive federal assistance. The Index of Medical Underservice (IMU) is the mechanism for determining MUA status. The IMU is considered by most researchers as a more sophisticated index than conventional population-to-physician ratios since it takes into consideration four factors:

- primary care population-to-physician ratio;
- infant mortality rate;
- proportion of the population aged 65 or over; and
- proportion of the population with an income below the poverty line.

The IMU score for an area is the sum of the weighted value for each factor. Values of the index range from 0 to 100, with lower scores indicating higher levels of medical underservice. Unlike other designations of underserviced areas, the IMU attempts to measure underservice by considering both medical care demand and supply (Lee, 1991; Office of Technology Assessment, 1990).

Other commonly used criteria include the ratio of 191 doctors per 100,000 population and other specialty-specific ratios recommended by GMENAC which used a needs-based approach to define the requirements (Kindig and Movassaghi, 1989). In Canada, the most often cited physician "requirements" are those recommended by the National Committee on Physician Manpower (Korcok and Geekie, 1976), the Federal/Provincial Advisory Committee on Health Manpower (1984) and the Royal College of Physicians and Surgeons of Canada (1988). In addition, various provincial ministries of health or task forces have come up with their own recommended population-to-physician ratios for medical workforce planning purposes.

The problem with identifying an area as overserviced or underserviced by comparing its population-to-physician ratio with an optimal ratio is that nobody seems to know how to objectively set an optimal ratio that accurately reflects local medical care needs. Many experts in physician workforce issues have proclaimed the futility of trying to define the optimal population-to-physician ratio.

> "...(T)here is no easy way - possibly no way - of gaining agreement among the experts, much less among the public, as to the criteria that should be used to assess whether a particular number of physicians is within the range of the social optimum. The answers fall in the complex realm of political economy in which each of many interest groups - the medical profession, the scientific-educational establishment, hospitals, health delivery organizations, federal and state legislators, consumer groups, and still other parties to the debate - has a perspective on the issue limited by their particular concerns and knowledge" (Ginzberg, 1983: p. 211).

> "It is important to state unequivocally at the outset that there is no technically correct or optimum number of physicians... In our view the past and continuing debate over the 'right' or 'optimum' number of physicians (or physician:population ratio) has not been constructive, has perpetuated the myth that better data or methodologies will resolve disagreements and has impeded a long overdue refocusing of attention on more fundamental and logically prior decision about the goals and structure of the health care system itself" (Stoddart and Barer, 1992: p. 698).

Ratios such as the Critical Medical Shortage Area ratio of 4,000:1 are mostly arbitrary in nature and established for administrative convenience with very little empirical basis. Even the use of widely endorsed

ratios, such as the GMENAC population-to-physician ratios, has been questioned. For example, it has been common practice to apply the GMENAC ratios to local areas. However, as Connor et al. (1994) have correctly observed, the nation was the unit of analysis in the GMENAC study, but these recommended national ratios are applied by some people to sub-national areas with no appropriate adjustments. In Canada, the National Ad Hoc Working Group on Physician Resource Planning (1995) has similarly warned against adopting national population-to-physician ratios to local use.

2.6.2 Comparing Physician Distributions in Different Areas

Another approach that has frequently been used to identify physician-shortage areas is to compare population-to-physician ratios of different areas. Areas that have a substantially lower ratio than others or that have ratios substantially below the national or provincial norm are often assumed to have insufficient physicians. Alternatively, some may conclude that maldistribution of physicians exists if areas exhibit widely divergent ratios. "Maldistribution", in this case, implies that the situation deviates from a normative pattern. Although such comparisons can be done by eyeballing the ratios, researchers have developed more sophisticated indices for this purpose.

While it is easy to distinguish perfect equality from inequality, when one level of inequality is used as a standard to assess another, different measures are needed. The Gini index and other indices of spatial distribution are often used to assess the difference between two distributions. (Politzer et al., 1998). It is useful to point out that the Gini index builds on the practitioner-to-population ratio. A Gini index is merely a descriptor of the relationship of two geographic distributions - that of the practitioners and that of the population (Morrow, 1977). The Gini index has been used to compare geographic distributions of physicians among regions or over time.

Another index is the location quotient which treats the ratio of physicians in a given geographic unit to the total number of physicians in all geographic units compared to the ratio of population in a given geographic unit to the total population of all geographic units. It signifies whether an area is underserviced or overserviced as compared to all areas. A value of 1.0 means that an area has exactly the number of physicians its population warrants, given its share of the total population. A value of less than 1.0 denotes an under-representation with respect to physicians while a value greater than 1.0 indicates a situation of over-representation. Anderson and Rosenberg (1990) have employed location quotients to study the distribution of physicians in Ontario over time. (A brief discussion of the methodological aspects of location quotient and Gini index can be found in Chapter 3).

Other researchers have used coefficients of variation to quantify inequalities with respect to geographic distribution of physicians. One commonly used approach is to compare the best-served region with the worst-served region within a province/state or country. Another approach, at the level of a province/state, is to compare the population-to-physician ratio in a rural area with that in an urban area (Blumenthal, 1994).

Whether the comparisons are between an observed population-to-physician ratio and an optimal ratio or between two or more observed ratios, the problem is the same, namely, an objective basis for determining physician requirement or underservice is often difficult to establish. Regional differences in physician supply, in and of themselves, do not provide sufficient grounds for deciding if there is an over- or under-supply of physicians, except in extreme situations. As Morrow (1977) has maintained, in order to move from the concept of "unevenness", as revealed by such indicators as the Gini index, to the concept of "maldistribution", many factors have to be considered. It is also necessary to apply some standard of adequacy to each health market area, a standard that takes into account the requirements of the population, morbidity levels, practitioner productivity, transportation, ability to pay, task delegation, and so forth.[7]

[7] It is interesting to note that in Ontario, the designation of a community as "underserviced" by the Underserviced Area Program of the Ministry of Health requires the assessment of several factors in addition to the population-to-physician ratio. These include population size and structure, financial impact analysis, previous recruitment efforts, socioeconomic status of the area, local demand for services and additional health service needs and resources (Ontario Ministry of Health, 1996).