# Importance of Capacity Assessment for an Early Staged-Research Network Designed to Eliminate Health Disparity: Lessons from RTRN

**Jae Eun Lee, DrPH**,
Research Centers in Minority Institutions Translational Research Network Data Technology Coordinating Center, Mississippi

**James Perkins, PhD**,
Research Centers in Minority Institutions Translational Research Network Data Technology Coordinating Center, Mississippi

**M. Edwina Barnett, MD, PhD, MBA**,
Research Centers in Minority Institutions Translational Research Network Data Technology Coordinating Center, Mississippi

**Daniel Sarpong, PhD**, and
Research Centers in Minority Institutions Translational Research Network Data Technology Coordinating Center, Mississippi

**Junghye Sung, MPH, ScD**
Department of Epidemiology & Biostatistics, School of Public Health, Jackson State University

## Abstract

**Background**—RCMI Translational Research Network (RTRN) is the first academic-based national network to address the problem of health disparities by integrating minority medical schools in a collaborative effort. While there was a great interest in forming the research network, limited systematic effort has been made in understanding members' existing capacity and future demand.

**Objective**—The aim of this study was to report the results of the RTRN Statistical Capacity Assessment and discuss the importance of an initial capacity assessment in building the biostatistical capacity of a research network in its early stage.

**Methods**—The assessment was based on survey responses submitted by program directors/ managers from 12 of the 18 RTRN institutions. In this assessment the capacity is defined as the statistical tools and human resources which are required for effective and efficient performance.

**Results**—A total of 52 biostatisticians (mean of 4.5 per site) were working for 12 RTRN institutions; 84% were fulltime employees, and 53% held a doctoral degree. On average, they had about 13 years of job experience. SAS, SPSS and STATA were the most frequently used and were selected as their major statistical software. A wide inter-institutional variability was found in number of biostatisticians (ranged from 1 to 8), mean years of experience in their position (4.5–20 years) and in major software (5–20), and the number of statistical software in use (1–11).

**Conclusion**—The initial capacity assessment provided valuable information on members' background and the network's research capacity which will be used as the basic data in developing

Address correspondence and reprint requests to Jae Eun Lee, DrPH; RTRN DTCC; Mississippi e-Center @JSU; 1230 Raymond Rd. Box 1800; Jackson, MS 39204; 601-979-0337; 601-979-0338 (fax); jae.lee@rtrn.net.

programs to build research capacity. Therefore, it is important to include the initial capacity survey and on-going evaluation of network activities when making business plans of research networks intended to reduce health disparities.

## Keywords

Research Network; Capacity Assessment; Statistical Capacity Building; RTRN; RCMI

## Introduction

Racial and ethnic minorities suffer from the highest rates of cardiovascular disease, cancer, diabetes, infant mortality, and other serious diseases.[1] Unfortunately, racial and ethnic minority groups have been traditionally less willing to participate in health research[2–4] and less informed of research opportunities.[5,6] Thus, minority groups remain underrepresented in many health research studies.[3,4] Much of the research performed on minorities can be more effective if the recruitment and research includes minority scientists, especially within the framework of a community setting.

In February 2002, a consortium of 18 NIH-funded Research Centers for Minority Institutions (RCMIs) convened to discuss creation of the RCMI Translational Research Network (RTRN). They envisioned a network that would conduct, for the first time, highly structured multi-site studies focusing on minority populations at minority institutions generating results that could be translated into major health benefits for all patients.

A variety of network models exist, which has led to a fundamental change in research culture.[7] However, RTRN is the first academic-based national network to specifically address the problem of health disparities across the continuum of basic-clinical community investigations through a network developed by integrating minority medical schools that have trained over 25 percent of the minority physician graduates in the United States. Indeed, by establishing a network of these schools including clinical and basic science disciplines, diverse racial and ethnic minority populations, and across vast geographic regions, RTRN represents the formulation of an innovative step toward understanding and reducing health disparities in minority populations. This network has established a framework for effective collaborations not only within RCMI institutions, but with larger research-intensive institutions and other organizations, allowing access to information and resources that move the entire research community toward a greater understanding of the tools needed to eradicate health disparities and equalize health care for all Americans.

While there is increasing interest and activity in forming research networks,[8] many lack a systematic effort to understand a network's existing capacity and future demands. Biostatisticians are key members in a research network since they help investigators across all aspects of research (ie, design, randomization, analysis, interpretation, and conclusion). Therefore, statistical capacity building may be an important issue in the early stage of development of a research network. To understand the background and skills of the biostatisticians in the RTRN and to develop possible strategies to maximize the network's statistical capacities, we conducted a basic assessment of the entire RCMI community by individually targeting each of the RCMI sites. The aim of this study is to report the results of the RTRN Statistical Capacity Assessment, identify the possible strategies to increase the RTRN statistical capabilities and discuss the importance of the initial capacity assessment in building biostatistical capacity for a research network.

## Methods

### Scope of the Assessment

In this assessment, statistical capacity is defined as the individual institution's and/or RCMI network's ability to engage in statistical functions required for all aspects of research (ie, developing protocols, designing case report forms, collecting and managing data, analyzing data and preparing manuscripts). Therefore, this assessment focused on the statistical resources currently available at the RCMI institutions and not on how the institutions operate them.

While recognizing the impact of the organizational system (such as the incentive system or accessibility to information) and psychological factors (such as morale or individual initiative) on the statistical capacity to undertake statistical activities, the assessment focused on the statistical tools and human resource capacities which enable biostatisticians to perform their functions effectively and efficiently. These are defined as: human resource capacity (the contract type, academic degree, and quantitative experience of the biostatistician); and statistical tool capacity (number of statistical software applications, the skill level of staff in utilizing the analytical software applications and the years of experience using the analytical tools).

### Assessment Procedure

The assessment relied on the information gathered in response to the 7 items dominating the survey: contract type, academic degree, software packages/programs in use, years of experience, current and upcoming statistical issues, and services needed.

Responses were provided by the program director/manager or a lead biostatistician at each institution. The respondents were encouraged to discuss the survey with others at their site to ensure thorough and accurate answers.

The survey form and memo were disseminated to the program directors/managers of the 18 RCMI institutions on April 7, 2008 and to ensure adequate participation two reminder notices were e-mailed on April 29 and May 27. Data collection closed on June 17, 2008. The completed assessment form was returned to the attention of the RTRN DTCC senior biostatistician via e-mail and fax. We obtained responses from 12 institutions.

### Analysis

Analyses were conducted at two levels, individual and institutional. Individual level analyses were conducted to determine the capacities of each individual biostatistician in the RTRN. For example, we estimated the average years of experience for all biostatisticians regardless of the parent institution and compared individual capacity with mean capacity of the network. By conducting the institutional level analyses, we determined if there were any differences in capacities among institutions. For example, we estimated mean years of experience for each institution and compared the averages among the institutions.

All statistical analyses for this assessment were carried out for descriptive purposes; therefore, no significance test was conducted.

## Results

### Distribution of Biostatisticians

The composition of biostatisticians was measured by determining the distribution of biostatisticians by the contract type and the academic degree. A total of 52 biostatisticians were

working for the 12 RCMI institutions that participated in the survey. Full-time biostatisticians were the most common contract type and 1 in 2 (53%) held a doctoral degree (Figure 1).

The mean number of biostatisticians per institution was 4.3 with a range from 1 to 8. Four institutions had 6 or more biostatisticians while 8 institutions had 5 or less, and only a single institution had one biostatistician. The ratio of doctoral to masters level biostatisticians also varied across institutions. Five institutions had a greater number of doctoral biostatisticians than masters biostatisticians; masters biostatisticians were greater in number in four of the institutions; and in two institutions there was an equal number of doctoral and masters biostatisticians. Four institutions had only doctoral biostatisticians while one institution did not have any biostatisticians with a doctoral degree.

### Experience

The experience of biostatisticians was measured quantitatively, not qualitatively. That is, they were asked how they had worked as a biostatistician. The biostatisticians in RCMI institutions had about 13 years of job experience associated with biostatistics (mean = 12.7 years; SD = 8.2 years). The doctoral-level biostatisticians had more experience than those with a masters degree. Approximately 3 in 5 biostatisticians (56%) had worked for more than 10 years. While 13% had 21 or more years of experience, 44% had less than 10 years of experience. Mean years of job experience were 9.4 (SD = 8.4) for the masters biostatisticians and 15.0 (SD = 7.2) for the doctoral biostatisticians (Figure 2).

Although variance in mean years of experience within an institute may not be a problem, variance across institutions may mean that well-specified programs should be prepared to reduce the gap between these institutions. Mean years of experience of each institution ranged from 4.5 to 20 years, meaning that with respect to biostatistical experience, both well- and less-experienced institutions co-existed in the RTRN.

### Statistical Analysis Tools in Use

A total of 22 statistical programs were identified as being in use at the RCMI institutions. Among them, SAS, SPSS and STATA were the most popular with 7 or more institutions using these programs. Specialized programs for power analysis and sample size determination (ie, PASS and Power and Precision) and for curve-fitting (ie, Graphpad, Graphpad Prism) were also used in a RTRN institution (Table 1). In addition, institutes associated with bioinformatics were using many different software packages for genetic analysis which were not included in this analysis.

The mean number of statistical applications in use was 5.5 per site (SD = 2.9). The number of statistical applications in use varied across institutions, ranging from 1 to 11. Half of the RTRN institutions (50%) were using at least 5 statistical package programs. Either SAS or SPSS was used in many of the RTRN institutions as their major statistical package program.

### Skill Levels in the Major Statistical Software Programs

Skill level in the statistical software was measured by determining if the biostatistician was using a statistical program that required high programming skill and how long the major statistical software had been in use. We asked respondents to indicate the major statistical software that their biostatistician(s) had used. A total of 11 different statistical software programs were indicated as their major software applications. Among them, SAS, SPSS and STATA were the most frequently selected by biostatisticians. About 88% of biostatisticians indicated SAS as their major analytical software, meaning that biostatisticians in RCMI institutions may be skillful in dealing with statistical software that requires programming skill (Table 2).

Mean years of experience in their major analytical software was 12.8 (SD = 7.6); about one in two biostatisticians were experienced in their major statistical software for longer than 10 years. About 1 in 10 biostatisticians had more than 20 years experience.

## Discussion

The basic assessment intended to describe the statistical resources of the RTRN institutions and not to determine if each institution or the whole network has adequate capacities to perform the statistical functions required in all aspects of research. The survey provided useful comprehensive information about the statistical personnel of the research network. The assessment suggested that there are several favorable signs in the statistical capacity of RTRN. That is, fulltime employment was a predominant contract type and more doctoral biostatisticians are working than master biostatisticians. Their job experience and skills in the major statistical software applications appear broad enough to perform the statistical activities needed for their current and upcoming projects. RTRN seems to possess appropriate statistical software packages for both generic and specialized analyses. However, unfavorable signs are a wide inter-institutional variability in the number of biostatisticians, experience and skill in major software, and statistical software they possessed.

Our survey data allowed us to identify the advantages and disadvantages of the network's statistical capacities and prepare capacity-building strategies to meet the particular needs described by the members. The biggest challenge in maximizing the statistical capacities of RTRN may be how to reduce the inter-institutional gap in the biostatistical capacities. The gap reduction may be achieved first by integrating the biostatistical capacities of all RCMI institutions and then sharing the knowledge and experience with institutions with fewer resources meaning that a whole system approach may be more appropriate. That is, the capacity of institutions with fewer resources can be increased by reinforcing the networking among member institutions rather than by focusing support on each individual member institute. Possible strategies to strengthen the networking could be a discussion bulletin board, LISTSERV for biostatisticians, providing password- protected server-based statistical software, planning a series of seminars, and formulating a RTRN biostatistics working group. The RTRN biostatistical working group could fill the gaps in training and experience by providing expertise for the entire network.

There were some limitations in interpreting the results of the survey. First, this is a partial assessment of sites' capacities. Although there are several factors (organizational and psychological) having an impact on statistical capacity, this assessment focused on the statistical tools and human resource capacities. Second, although more than half of the 18 RTRN institutions responded to the survey, a significant number of institutions did not participate. The results, therefore, may not be an accurate accounting of the network's total capacity. Third, we relied on the responses provided by the program director/manager or a lead biostatistician at each institution, not by the individual biostatisticians– this may lead inaccurate responses. However, we encouraged the respondents to discuss the survey with relevant staff at their site to ensure thorough and accurate answers. Lastly, in this survey, we did not use a well-structured qualitative measure in assessing the biostatisticians' ability and depth of knowledge and research skills. A future survey using well, predetermined measures targeting individual biostatisticians is required.

Despite these limitations, our survey provided valuable information on member biostatisticians' background and the network's research capacity which will be used as the basic data in developing programs to build research capacity. It is, therefore, important to include the initial capacity survey and on-going evaluation of network activities when building the business plan of research network.

# References

1. Institute of Medicine Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington, DC: National Academy Press; 2003.

2. Shavers VL, Lynch CF, Burmeister LF. Racial differences in factors that influence the willingness to participate in medical research studies. Ann Epidemiol 2002;12(4):248–256. [PubMed: 11988413]

3. Ford BM, Evans JS, Stoffel EM, Balmana J, Regan MM, Syngal S. Factors predicting willingness to participate in cancer genetic epidemiologic research. Cancer Epidemiol Biomarkers Prev 2006;15(7): 1355–1359. [PubMed: 16835336]

3. Ford BM, Evans JS, Stoffel EM, Balmana J, Regan MM, Syngal S. Factors associated with enrollment in cancer genetics research. Cancer Epidemiol Biomarkers Prev 2006;15(7):1355–1359. [PubMed: 16835336]

4. Brown DR, Topcu M. Willingness to participate in clinical treatment research among older African Americans and Whites. Gerontologist 2003;43(1):62–72. [PubMed: 12604747]

5. Slomka J, Ratliff EA, McCurdy SA, Timpson A, Williams ML. Decisions to participate in research: views of underserved minority drug users with or at risk for HIV. AIDS Care 2008;20(10):1224–1232. [PubMed: 18608070]

6. Wendler D, Kington R, Madans J, et al. Are racial and ethnic minorities less willing to participate in health research. PLoS Med 2006;3(2):e19, 201–210. [PubMed: 16318411]

7. Ried K, Farmer EA, Weston KM. Setting directions for capacity building in primary health care: a survey of a research network. BMC Fam Pract 2006 Feb 9;7:8. [PubMed: 16466583]

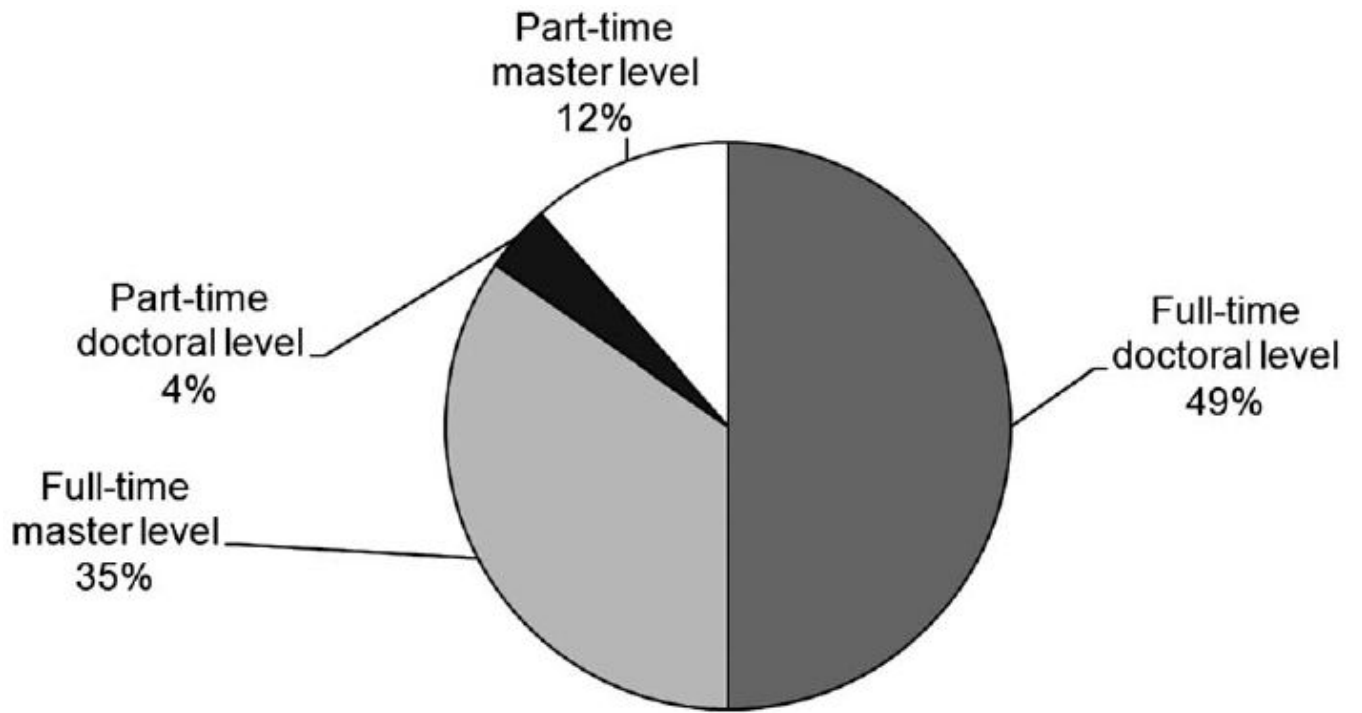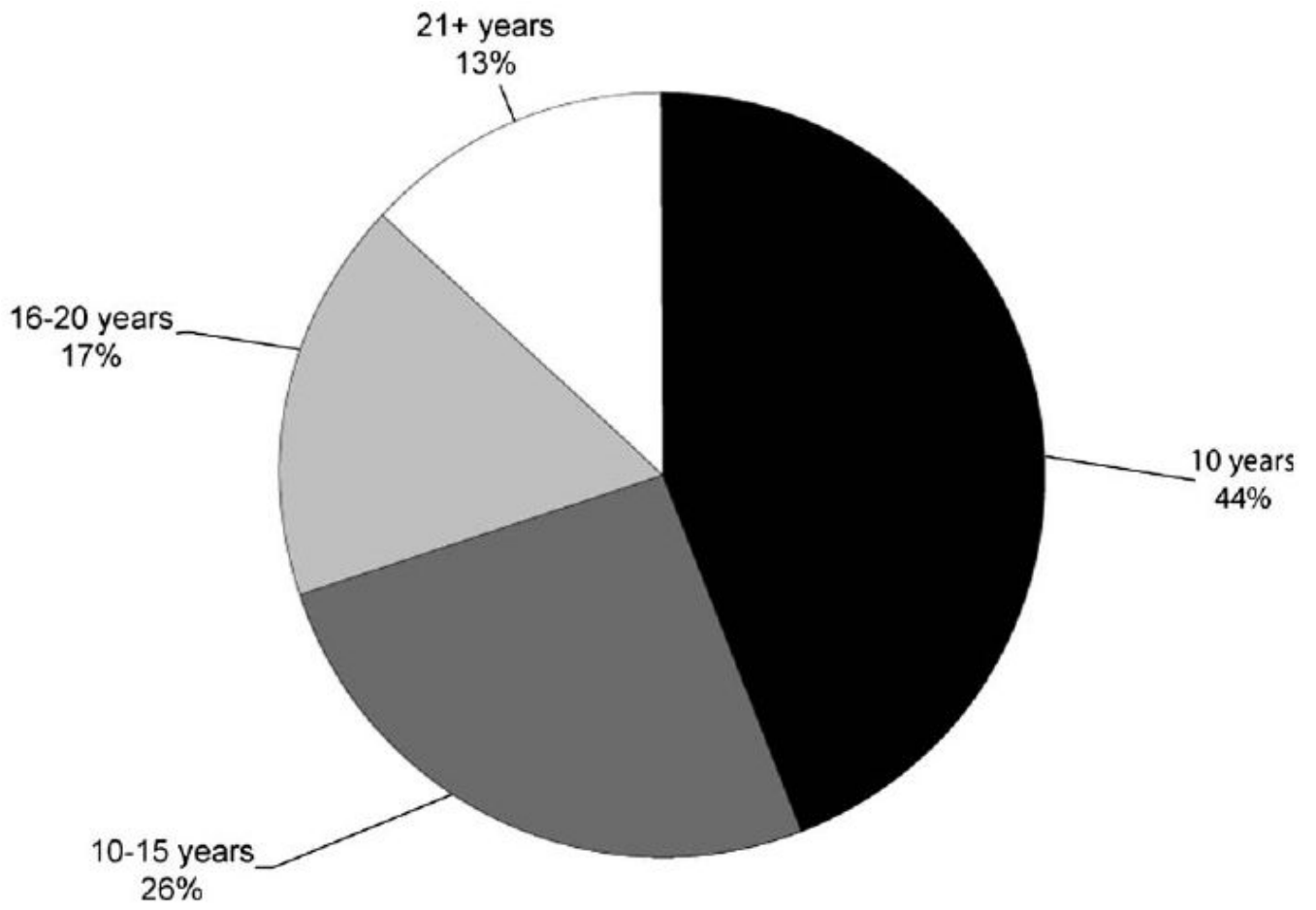8. Gunn JM. Should Australia develop primary care research networks? Med J Aust 2002;177(2):63–66. [PubMed: 12098340]

**Fig 1. Distribution of biostatisticians by academic degree and contract type**

**Fig 2. Distribution of biostatisticians by years of biostatistical experience**

**Table 1**

**Statistical software in use**

| Name of the software | Institutions | |
|---|---|---|
| | Number | % |
| SAS | 11 | 91.7 |
| SPSS | 10 | 83.3 |
| STATA | 7 | 58.3 |
| SUDAAN | 5 | 41.7 |
| EPINFO | 4 | 33.3 |
| Mathematica | 3 | 33.3 |
| Matlab | 3 | 25.0 |
| S-Plus | 3 | 25.0 |
| Minitab | 3 | 25.0 |
| R | 2 | 16.7 |
| ActivStats, BMDP, GraphPad Instat version 3.06, Graphpad Prism, Mathematica, PASS, Power and Precision from Biostat (v. 2.0.4), Prism from GraphPad (v. 4.03), Stat View, Systat, @RISK, Neuroshell2 | 1 | 8.3 |

**Table 2**
**Major statistical software being used by biostatisticians**

| Major software in use | % |
| --- | --- |
| SAS | 87.5 |
| SPSS | 60.0 |
| STATA | 40.0 |
| Epiinfo | 15.0 |
| R | 15.0 |
| S-PLUS | 12.5 |
| SUDAAN | 12.5 |
| Matlab | 10.0 |
| MINITAB | 10.0 |
| Statistics Toolbox™ | 2.5 |
| StatView | 2.5 |