

Center for Regulatory Effectiveness

Suite 700
11 Dupont Circle, N.W.
Washington, DC 20036-1231
Tel: (202) 265-2383 Fax: (202) 939-6969
secretary1@mbsdc.com www.TheCRE.com

May 1, 2007

Mr. Richard A. Spires
Chief Information Officer
Internal Revenue Service
U.S. Department of the Treasury
1111 Constitution Ave., NW -- Rm. 3137
Washington, DC 20224

Re: Data Quality Act issues in evaluating the current tax exemption for non-profit hospitals

Dear Mr. Spires:

It is our understanding that the IRS has been considering whether there should be any changes to the rules or guidance covering the current federal tax exemption for nonprofit hospitals. During the last few years there have been Congressional hearings and a number of government studies on this subject.

The Center for Regulatory Effectiveness (CRE) has devoted considerable attention to this and other health care issues, including the application of epidemiologic analysis to rulemaking on public health issues. CRE has also been a leader in enactment and implementation of the Data Quality Act (Sec. 515 of Public Law 106-554, 44 U.S.C. § 3516, note) and the OMB and agency-specific implementing guidelines. In view of the potential for IRS action and dissemination of public information on this issue -- whether in the form of rulemaking, guidance, Congressional testimony, or any other type of public information -- we wish to alert you to some of the key Data Quality issues that it appears might be involved in any IRS analysis of this subject.

All of our comments concern issues arising under the OMB and IRS Data Quality guidance. At the outset, a key point we wish to make in this regard is that, although IRS has not disseminated any information on the issues we discuss below, if it were to do so, it would be subject to the Data Quality guidance provisions that apply to agency reliance on non-agency data. If, in its own information dissemination, the IRS were to rely on non-agency data, those data would have to meet the same Data Quality standards that apply to dissemination of federal agency information. Therefore, the main purpose of this letter is to advise you of Data Quality

Center for Regulatory Effectiveness

problems we have observed regarding certain published studies that are often cited in support of the current federal tax exemption for nonprofit hospitals that the IRS might consider incorporating into its own analysis of the issues. Similar problems can be observed with other studies concerning this subject, but the studies discussed herein provide good examples of several important Data Quality deficiencies.

We do not have a policy position on the issue of the tax exemption; we are concerned only that any IRS give careful scrutiny to data claims being made by participants in the debate pursuant to its Data Quality pre-dissemination review authority and responsibilities.

The three published studies we will use as examples to comment on from a Data Quality perspective are --

Schlesinger M and Gray BH. 2006. How nonprofits matter in American medicine, and what to do about it. *Health Affairs* 25(4):W287-303. Epub June 20, 2006. (Hereafter referred to as "Schlesinger/Gray 2006" or "S/G".)

Devereaux PJ *et al.* 2002. A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals. *CMAJ* 166(11):1399-1406. (Hereafter referred to as "Devereaux et al. 2002".)

Devereaux PJ et al. 2004. Payments for care at private for-profit and private not-for-profit hospitals: A systematic review and meta-analysis. *CMAJ* 170(12):1817-24. (Hereafter referred to as "Devereaux et al. 2004".)

These three articles are cited frequently by proponents of the current nonprofit tax-exemption. The arguments put forth are that the current tax exemption for non-profit hospitals is justified not only by the community benefits they provide in the form of "charity care", but also by higher quality care and lower charges compared to for-profits. They argue that even though the level of uncompensated care provided might not sufficiently justify the tax-exemption taken alone, when taken in the aggregate, charity care, higher quality care, and less expensive care do justify the current exemption.¹

As explained below, however, these three studies, and ones like them, appear significantly flawed and non-compliant with the federal Data Quality standards, and even their analytical results are exaggerated, both in the reports themselves, and by those who rely on them.

First, however, it is necessary to provide a brief discussion of the Data Quality standards that would apply if the IRS were to rely on any of these studies.

¹ S/G at W296.

Applicable Requirements of the OMB and IRS Data Quality Guidelines

Agency (IRS) reliance on non-agency data

Although the DQA guidelines issued by OMB (government-wide) and IRS (agency-specific) apply only to agency information products disseminated to the public², the OMB government-wide guidance states: "[I]f an agency, as an institution, disseminates information prepared by an outside party in a manner that reasonably suggests that the agency agrees with the information, this appearance of having the information represent agency views makes agency dissemination of the information subject to these guidelines."³ Similarly, "reliance" by an agency on a non-agency study would constitute agency dissemination of that study.⁴ In such cases, the outside data relied upon by the agency must meet the same standards ("quality, objectivity, utility, and integrity", as defined) as if they were agency data, including the "capable of being substantially reproduced" standard for "influential scientific, financial, or statistical information."⁵

The IRS guidance does not specifically address the issue of reliance on non-agency data; it only addresses IRS "sponsorship" of information disseminations -- *i.e.*, situations where IRS "directs a third party to distribute information or the IRS has the authority to review and approve the information before release."⁶ Nevertheless, the OMB guidance is controlling -- all agencies' guidance must be consistent with the OMB guidance.⁷

² The IRS Data Quality guidelines can be accessed at <http://www.irs.gov/pub/irs-utl/infoqualityguidelines.pdf>. The guidelines state, in their "Scope" section, that they exempt "[t]estimony of officials, information or drafting assistance provided to Congress related to pending or proposed legislation that has not been previously disseminated to the public in another manner." At p. 2. Such an exemption is not authorized by the Act or the OMB government-wide guidelines and is contrary to OMB's June 10, 2002, memorandum to agencies commenting on issues regarding drafting of their agency-specific guidelines (available at http://www.whitehouse.gov/omb/inforeg/iqg_comments.pdf). However, this IRS statement of exemption is later contradicted in the IRS guidelines at p. 13, where the guidance states that it applies to "reports to Congress". Moreover, such an exemption is not set out in the "definitions" section of the IRS guidance with regard to the meanings of "information" or "dissemination", where other exemptions are specified.

³ 67 FR 8452, 8454 1st col., Feb. 22, 2002.

⁴ *Id.* at 8457 2d and 3d cols. See also the June 10, 2002 OMB/OIRA memorandum to the President's Management Council at 6-7 (available at http://www.whitehouse.gov/omb/inforeg/iqg_comments.pdf).

⁵ *Id.* and 8460. Such information is "influential" if "the agency can reasonably determine that dissemination of the information will have or does have a clear and substantial impact on important public policies or important private sector decisions." *Id.* at 8457 3d col.

⁶ Fn. 2 at 2.

⁷ 44 U.S.C. § 3506(a)(1)(B).

Center for Regulatory Effectiveness

The quality standards

The most basic quality standard that all agency-disseminated information must meet is that of "objectivity", which is defined to mean whether disseminated information is being presented in an "accurate, clear, complete, and unbiased manner." "In a scientific, financial, or statistical context, [objectivity requires that] the original and supporting data shall be generated, and the analytic results shall be developed, using sound statistical and research methods."⁸ The original government-wide OMB guidelines of Feb. 22, 2002 state that journal peer review raises a presumption that the published data is objective; however, that presumption can be rebutted by a persuasive showing of non-compliance. Additionally, the original government-wide guidance has since been supplemented by OMB government-wide guidance on peer review that, as explained below, contains additional requirements for agency peer review of "influential scientific information" and "highly influential scientific assessments".

The IRS guidelines do not mention peer review, particularly in their definition of "objectivity", although they do contain numerous references to internal IRS quality reviews (*e.g.*, in the section of "objectivity" under "Operating Divisions"). Nevertheless, the OMB peer review guidelines are binding on the IRS.⁹

Peer review requirements for agency scientific information and assessments

In January 2005, OMB finalized guidelines with extensive peer review requirements for "scientific information" disseminated by agencies.¹⁰

The guidance sets out requirements for peer review of "influential scientific information", and more rigorous peer review requirements for "highly influential scientific assessments". Basically, the peer review process for "highly influential scientific assessments" requires external peer review, more public participation in both peer review planning and execution, and greater transparency (*e.g.*, a peer review report and agency responses to significant criticisms). In addition, the guidance sets out requirements for agencies to publish on the Internet their plans for peer review of both "influential scientific information" and "highly influential scientific assessments" and to accept public comments on those plans.¹¹

The OMB peer review guidelines define "scientific information" very broadly to mean "factual inputs, data, models, analyses, technical information, or scientific assessments related to such disciplines as the behavioral and social sciences, public health and medical sciences, life and earth sciences, engineering, or physical sciences. This includes any communication or

⁸ 67 FR 8452, 8459-60, Feb. 22, 2002 and Appendix A of the IRS guidelines, fn. 2.

⁹ Fn. 7, above.

¹⁰ 70 FR 2664, Jan. 14, 2005.

¹¹ *Id.* at 2676-77.

Center for Regulatory Effectiveness

representation of knowledge such as facts or data, in any medium or form including textual, numerical, graphic, cartographic, narrative, or audiovisual forms."¹²

The term "influential scientific information" is defined as meaning "scientific information the agency reasonably can determine will have or does have a clear and substantial impact on important public policies or private sector decisions"¹³

A "highly influential scientific assessment" is defined as meaning a "scientific assessment" that "(i) Could have a potential impact of more than \$500 million in any year, or (ii) Is novel, controversial, or precedent-setting or has significant interagency interest."¹⁴

A "scientific assessment" (as distinguished from "scientific information") is defined as meaning "an evaluation of a body of scientific or technical knowledge, which typically synthesizes multiple factual inputs, data, models, assumptions, and/or applies best professional judgment to bridge uncertainties in the available information. These assessments include, but are not limited to, state-of-science reports; technology assessments; weight-of-evidence analyses, meta-analyses"¹⁵

The preamble to the OMB peer review guidelines makes clear that journal peer review does not suffice to meet the Data Quality peer review requirements, nor does prior peer review raise the same presumption of objectivity for agency peer review purposes as in the case of agency information that is not of a type covered by the peer review guidelines (influential scientific information or highly influential scientific assessments). The preamble states with regard to "influential scientific information":

Publication in a refereed scientific journal may mean that adequate peer review has been performed. However, the intensity of peer review is highly variable across journals. There will be cases in which an agency determines that a more rigorous or transparent review process is necessary. For instance, an agency may determine a particular journal review process did not address questions (*e.g.*, the extent of uncertainty inherent in a finding) that the agency determines should be addressed before disseminating that information. As such, prior peer review and publication is not by itself sufficient grounds for determining that no further review is necessary.¹⁶

¹² 70 FR 2664, 2667 2d and 3d col. and 2675 1st col. (Jan. 14, 2005).

¹³ 70 FR at 2675 1st col.

¹⁴ 70 FR at 2675 3d col.

¹⁵ 70 FR At 2675 1st col.

¹⁶ 70 FR at 2671 1st col.

Center for Regulatory Effectiveness

When any peer review is conducted, the peer reviewers must apply the OMB and agency Data Quality standards.

The peer review requirements for "highly influential scientific assessments" are not qualified by any statements regarding the significance of prior agency or journal peer review, and the requirements (*e.g.*, public participation and transparency) are clearly beyond what is involved in internal agency or journal peer review

An IRS assessment of "community benefits" and justification of the tax exemption for hospital nonprofits that involved the articles of concern here would certainly be "influential scientific information" and would very likely be a "highly influential scientific assessment" requiring public participation, external peer review and peer review planning.

Agency pre-dissemination review

While peer review is obviously one type of pre-dissemination review, the original OMB government-wide guidelines also set out broad requirements for agency pre-dissemination review that preceded the peer review guidance and are still in effect. The 2002 OMB guidelines state:

As a matter of good and effective agency information resources management, agencies shall develop a process for reviewing the quality (including the objectivity, utility, and integrity) of information before it is disseminated. Agencies shall treat information quality as integral to every step of an agency's development of information, including creation, collection, maintenance, and dissemination. This process shall enable the agency to substantiate the quality of the information it has disseminated through documentation or other means appropriate to the information.¹⁷

The IRS guidelines contain numerous references to internal IRS quality reviews (*e.g.*, in the section of "objectivity" under "Operating Divisions").

Data Quality Deficiencies in the Three Articles

1. Schlesinger/Gray 2006

This article relies primarily on a type of "meta-analysis" (if it can even be justifiably called that) sometimes referred to as "vote counting" that is considered unreliable and therefore lacking "objectivity" within the meaning of the Data Quality guidance of both OMB and IRS. This "vote counting" is presented mainly in its Exhibit 2, which simply counts up the number of positive, negative, and non-statistically-significant studies on various attributes of for-profit and nonprofit hospitals (and nursing homes). "Vote counting" is definitely not regarded as a reliable methodology by the scientific community, since, as in this article, it does not take into account study quality, effect size, or confidence intervals, or attempt to describe or quantify uncertainty.

¹⁷ 67 FR at 8459 1st col.

Center for Regulatory Effectiveness

There can be a substantial subjective element in vote counting, since it requires the vote counter to at least determine the relevance of each of the studies -- an aspect that is very important in this case because, as will be shown, the subjective element introduces substantial bias by counting as relevant studies favoring the current non-profit tax-exemption that appear to be either clearly not relevant or of very arguable or marginal relevance.

Also, as discussed later on below, the S/G article is biased in citing the two Devereaux et al. study reports as if they were authoritative, when in fact they are problematic and are more likely to indicate a lack of differences between for-profits and nonprofits.

This is what the Cochrane Collaboration -- the prestigious independent non-profit medical research group -- has to say about vote counting in the COCHRANE HANDBOOK FOR SYSTEMATIC REVIEWS OF INTERVENTIONS:

8.6.10 Use of vote counting for meta-analysis

Occasionally meta-analyses use “vote-counting” to compare the number of positive studies with the number of negative studies. Vote-counting is limited to answering the simple question “is there any evidence of an effect?” Two problems can occur with vote-counting, **which suggest that it should be avoided whenever possible**. Firstly, problems occur if subjective decisions or statistical significance are used to define “positive“ and “negative” studies (Cooper 1980, Antman 1992). . . . Secondly, vote-counting takes no account of the differential weights given to each study. Vote-counting might be considered as a last resort in situations when standard meta-analytical methods cannot be applied (such as when there is no consistent outcome measure).¹⁸

Because vote counting is "unreliable", it violates the DQ "objectivity" standard. It also is not objective because it should, but does not, disclose "error sources affecting data quality", and does not use "sound statistical and research methods". In particular, it makes no effort to explore the great and obvious heterogeneity among the studies tabulated in its Exhibit 2.¹⁹

Since Exhibit 2 of the article contains the vote counting and is central to the article's theme and conclusions, it is useful to examine some selected portions of Exhibit 2 for the quality and relevance of the studies that the authors have selected to support certain vote counts. This

¹⁸ Version 4.2.6, September 2006 (available at <http://www.cochrane.org/resources/handbook>) (emphasis added). Other reference works could be cited for similar views on vote counting as an analytical method.

¹⁹ The specific references supporting Exhibit 2 are not contained in the published article; instead they are made available on the journal's website as supplementary material in the form of a Technical Appendix. The observations below are based on those references and on abstracts of the cited studies, not the full studies. Exhibit 2 gives an Internet address for accessing the comprehensive list of studies supporting the Exhibit, but the address is not operative. Nevertheless, the list of studies can be found through the Health Affairs website by obtaining information on the article from its archives. We note that many of the references in the Technical Appendix for the studies discussed below are incomplete (lacking dates, journal title, etc.) or are inaccurate.

Center for Regulatory Effectiveness

examination illustrates the unreliability of the vote counting methodology as the basis for the types of conclusions asserted by the authors.

With regard to hospitals, the authors draw three conclusions in particular about nonprofit vs. for-profit hospitals from Exhibit 2, as well as from the two Devereaux et al. studies discussed below: These three conclusions are (at 296) that nonprofit hospitals (1) charge patients less; (2) have fewer adverse outcomes; and (3) provide more uncompensated care.

Initially, it should be observed from Exhibit 2 that for almost all categories there are a substantial number of studies indicated as supporting no significant difference between nonprofit and for-profit hospitals; however, some sub-categories indicate a substantially greater number of studies supporting a nonprofit advantage compared to those supporting a for-profit advantage. This is apparently the basis for the S/G assertions that non-profits are superior in those categories. Below we examine whether it appears that the individual studies cited as supporting the vote counts given in favor of nonprofits in certain key categories and sub-categories are in fact relevant and supportive.

Charges to patients: The most pertinent sub-category here is "Revenues/Charges Per Admission," and Exhibit 2 shows that there are nine studies supporting a non-profit advantage, four supporting "no significant differences"²⁰, and none supporting a for-profit advantage.

The nine studies asserted to support a non-profit advantage are referenced in footnote 14 of the Technical Appendix. Abstracts for eight of the nine studies were obtained. One of the studies is Devereaux et al. 2004, which is discussed below as only very barely and questionably qualifying as a study supporting non-profits, and of questionable value. The study by Clement and Grazer concerned the effects of HMO penetration on revenues, expenses, and operating margins of all types of hospitals, and it was not aimed at comparing charges to patients by nonprofits with for-profits. The Younis et al. study was aimed at identifying the determinants of hospital profitability among all types of hospitals. The Bazzoli et al. study examined the financial performance of hospitals belonging to health networks and systems, and did not compare charges by nonprofits with charges by for-profits. The Melnick et al. study examined the impact of mergers and acquisitions on price competition and charges "regardless of ownership status". The Connor et al. study examined the price impacts of mergers and HMO penetration on all types of hospitals. The Meurer et al study looked at charges for only a single health endpoint, asthma. The Lynk study examined the impact of mergers on efficiency and operating costs. The Eskoz and Peddecord study examined charges at hospitals only in California during a period prior to 1985, a period when the hospital industry environment was very different from today.

In all, it appears that perhaps only two of the nine studies that purportedly support a non-profit advantage with regard to charges are clearly relevant or provide slim and limited support.

²⁰ Presumably "no significant differences" refers to lack of statistical significance rather than lack of a showing of a significant level of statistically significant difference. Otherwise, the two Devereaux et al. articles -- which show extremely minor differences, although statistically significant using their methodology -- would not be referenced as firmly supporting a nonprofit advantage.

Center for Regulatory Effectiveness

Adverse outcomes: On this subject, two sub-categories of references stand out as appearing to provide a distinct advantage to non-profits: "Mortality rates (post discharge)" and "Process measures of quality." For post-discharge mortality rates, Exhibit 2 shows seven studies supporting a non-profit advantage, nine studies showing no significant differences, and only one study supporting a for-profit advantage. For process measures of quality, Exhibit 2 shows 5 studies as supporting a non-profit advantage, one study as indicating no significant differences, and no studies supporting a for-profit advantage.

Mortality rates (post discharge)

The Shen study addressed only one disease endpoint, acute myocardial infarction, and found only a 3-4% difference between for-profits and nonprofits. In addition, as discussed below, the Congressional Budget Office study of 2006²¹ found that on average non-profits are located in areas with better socioeconomic conditions, which could easily account for such a small difference in outcomes.

The Thomas et al. study focused only on hospitals in Utah and Colorado in 1992, and the difference favoring non-profits was barely statistically significant at a lower confidence interval of 1.03.

The Gowrisankaran and Town study examined only pneumonia outcomes in Southern California from 1989 to 1994, and focused mainly on the difference that adjusting for severity of illness had on the results. As noted elsewhere herein, locational and related socioeconomic conditions are likely to affect severity of illness and outcomes.

The al-Haidar and Wan study is misclassified in the Exhibit: It examined only in-hospital mortality, not post-discharge mortality. This study also, although published in 1991, utilized 1984 Medicare data, which immediately makes it suspect as to relevancy because the Medicare prospective payment service program was phased in during 1984, and during and after the phase-in period there were significant changes in the hospital operating environment. Additionally, the whole industry structure has changed radically since 1984, and, as pointed out below in connection with the Devereaux et al. 2002 study, the recent Congressional Budget Office study of 2006 found that on average non-profits operate in a better socioeconomic environment, which undoubtedly results in advantages with regard to outcomes. In this connection, the study specifically notes that it was not able to control for severity of illness.

The Hartz et al. study utilized 1986 data, and therefore is suspect as to current relevancy due to subsequent and ongoing changes in the hospital industry and the proximity to the 1984 changes in Medicare.

In summary, five of the seven studies shown in Exhibit 2 as supporting a non-profit advantage with regard to post-discharge mortality are of questionable relevance or provide only very marginal support.

²¹ Fn. 28, below.

Center for Regulatory Effectiveness

Adverse Outcomes (other than mortality)

Although this subcategory does not show a distinct advantage for non-profits, we note that although the Exhibit indicates five studies supporting a non-profit advantage, the Technical Appendix provides only three references, and two of those are doctoral theses rather than journal articles.

Process Measures of Quality

The Jha et al. study found a small but significant advantage for nonprofits, but it also found that a factor influencing an increase in performance was size and location in the Northeast or Midwest, a factor favoring nonprofits.

The Sloan et al. study found that for-profit hospitals had lower rates for use of aspirin and ACE inhibitors (although barely significant), but that they had higher rates for use of higher-tech treatments such as PTCA (percutaneous transluminal coronary angioplasty) and CABG (coronary artery bypass grafts). The authors concluded that the differences in treatment measures did not affect outcomes, and appeared to reflect demand for higher-tech treatments at for-profit hospitals.

The Weinstein study is not a published article but a doctoral thesis, and deals only with cesarean deliveries in California. Moreover, a recent (2007) study concluded that cesarean delivery is not a reliable measure of quality.²²

The Keeler et al. study used data from 1981 to 1982 and 1985 to 1986 to examine whether the Medicare prospective payments system (phased in from Oct. 1983 to Sept. 1984) affected mortality for certain disease endpoints. The study found that there were large variations in outcomes from State to State, but with teaching, larger, and urban hospitals having an advantage. The study also found that small, non-teaching hospitals were closing the gap between 1981 and 1986.

The Placek et al. study examined the changes in rates of cesarean deliveries between 1965 and 1981. It does not appear to have any relevance to current conditions, and cesarean delivery as a measure of quality of care is not generally accepted.²³

In summary, only one of the five studies cited as supporting a non-profit advantage for quality measures arguably supports such an advantage (Jha et al.), and that advantage is small and dependent on locational attributes.

²² Bailit JL. 2007. Measuring the quality of inpatient obstetrical care. *Obstet Gynecol Surv* 62(3):207-13.

²³ *Id.*, and see Hospital Quality Alliance measures specified by JCAHO and CMS (at http://www.cms.hhs.gov/HospitalQualityInits/downloads/HospitalHQA2004_2007200512.pdf, note).

Center for Regulatory Effectiveness

Uncompensated care: Exhibit 2 cites five studies as showing a nonprofit advantage for locating in low-income areas, with no studies showing no significant difference or favoring for-profits. It also shows 12 studies supporting a non-profit advantage for treating uninsured patients vs. 6 studies showing no significant difference, and no studies showing a for-profit advantage.

Locating in low-income areas

The Norton and Staiger study found that when for-profit and nonprofit hospitals are located in the same area, they serve an equivalent number of uninsured patients; but that for-profit hospitals tend to locate more in better-insured areas. The finding on locational preference of for-profits is at odds with the 2006 Congressional Budget Office study, discussed below, which is not cited.

The Homer et al. study is a 1984 study of nursing homes, and it should be cited in the Exhibit section on nursing homes rather than hospitals. Additionally, it is of very doubtful relevance in view of its age.

The Mullner and Hadley study is a study in growth of market share of chain-operated proprietary hospitals during the period 1973-1984. It does not appear relevant either with regard to substance or the time period covered.

The Kushman and Nuckton study is even older -- from 1977 -- and it is puzzling why it is considered relevant, since it studied the responsiveness of proprietary hospital bed numbers to population increases and found that the for-profits were more responsive than nonprofits.

The Congressional Budget Office study²⁴ of nonprofits and community benefits that was released December 6, 2006, after the S/G article was published, contains findings supporting a for-profit advantage with regard to locating in low-income areas. The CBO study focused on five States with the best data, and it found that in four of those States nonprofits were found to be more likely than for-profits to be located in counties with either lower poverty rates, lower rates of uninsurance, or both. It also found that "[a]t the national level, nonprofit hospitals, compared with for-profit hospitals, tend to be located in higher-income counties. . . ." At 16 and n. 48.

In summary, at least three of the five studies cited as supporting a nonprofit advantage for locating in low-income areas appear to be outdated or irrelevant, and the other two are contradicted by the more recent CBO study.

Treating uninsured patients

Exhibit 2 cites 12 studies as supporting a nonprofit advantage, six studies as showing no significant difference, and no studies supporting a for-profit advantage.

Three of the cited studies (Wolff and Schlesinger, Schlesinger et al., and Olfson and Mechanic) are restricted to psychiatric hospitals or psychiatric care, and these are specialized

²⁴ Fn. 28, below.

Center for Regulatory Effectiveness

types of care and facilities that should not be compared with general acute-care hospitals. In addition, the Wolff and Schlesinger study is very outdated, having used data from 1975, 1980, and 1986.

Three of the other cited studies were limited to California (Clement et al., Campbell and Ahern, and Seidman and Pollack). In 2005, the Government Accountability Office produced a study in which it found that for-profits provided more uncompensated care than nonprofits in California.²⁵ The GAO study is not cited in Exhibit 2.

The Frank et al. study is clearly outdated, having used data from 1979-84.

In summary, at least seven of the 12 studies asserted to show a nonprofit advantage are either not relevant or outdated.²⁶

Any commentary on the accuracy of the S/G article with regard to uncompensated care (which S/G refer to as "treating uninsured patients" -- not necessarily the same) would be incomplete without comparing its conclusions with those contained in the 2005 GAO²⁷ and 2006 CBO²⁸ studies on uncompensated care by non-profits.

The information in the Schlesinger/Gray 2006 article on uncompensated care is at variance with the GAO (2005) and CBO reports (2006). As noted before, the GAO report was published prior to the S/G article, but the CBO report was published after. Both the GAO and CBO reports employ more sophisticated and reliable methodology than the S/G article. The S/G article uses simplistic vote counting, whereas the GAO and GBO reports use statistical distributions in a transparent manner. The S/G and GAO/CBO studies report very different results. The S/G report gives the appearance of nonprofits having a distinct advantage with regard to uncompensated care; however, the GAO and CBO data show that the difference is relatively small, very unevenly distributed (*i.e.*, many non-profits do not provide more uncompensated care than for-profits), and, in the case of the CBO report, not commensurate to any extent with the tax benefits provided by the nonprofit tax exemption. The CBO study reported:

The difference in the total amount of uncompensated care provided by nonprofit and for-profit hospitals is largely attributable to the fact that nonprofit hospitals account for a much larger share of the hospital market than did for-profits. . . .

²⁵ Fn. 27, below.

²⁶ The Marmot et al. citation is for a "handbook" published in 1987. There was no abstract available for the Eckels and Miller study, published in 1988. Both are of questionable relevance because of their study periods.

²⁷ Government Accountability Office. 2005 Nonprofit, for-profit, and government hospitals, uncompensated care and other community benefits. GAO-05-743T.

²⁸ Congressional Budget Office. 2006. Nonprofit hospitals and the provision of community benefits. 2006. CBO Pub. No 2707, Dec. 2006.

Center for Regulatory Effectiveness

When regression techniques were used to adjust for the hospitals' size and location and for the characteristics of the local populations, nonprofit hospitals were estimated to have an average uncompensated-care share that was 0.6 percentage points higher than that for otherwise similar for-profit hospitals. That estimated difference corresponds to nonprofit hospitals in the five selected states providing between \$100 million and \$700 million more in uncompensated care than would have been provided if they had been for-profits.²⁹

In connection with this care cost estimate, CBO also noted: "JCT [the Congressional Joint Committee on Taxation] also estimated the value of some of the tax exemptions for nonprofit hospitals located in the five states In those five states, the exemptions from federal and state corporate income taxes, state and local sales taxes, and local property taxes were valued at \$2.5 billion. (Two important categories of tax exemptions -- tax-exempt bond financing and the deductibility of charitable contributions -- were included in the national totals but were not available for the five states and are not included in the five-state total.)" (At 3) Thus, nonprofit tax advantages greatly outweighed the cost of providing uncompensated care.³⁰

The CBO report on the use of tax-exempt bonds by nonprofit hospitals estimated lost government tax revenues at approximately \$1.8 billion. Also of interest are the report's comments on studies of differences between nonprofits and for-profits:

Various analyses have thus examined not only differences in for-profit and nonprofit hospitals' "outputs," such as uncompensated care and quality of care, but also differences in their operational measures such as cost, profitability, pricing, adoption of technology, and financial policies.

Although those analyses find some differences between for-profit and nonprofit hospitals, in some locations, for some services, they do not reveal consistent patterns.⁵ Consequently, analysts and policymakers have reached different conclusions about whether the level of community benefits provided by nonprofit hospitals justifies the tax preferences they receive."³¹

Any analysis that might be produced by the IRS on this issue could not be considered "complete" (and therefore objective) unless it included the CBO report and contrasted the methodologies it used with the simplistic (and biased) vote counting used in the S/G article.

²⁹ *Id.* at 2.

³⁰ The Congressional Budget Office has reported that the total value of federal tax exemptions provided to nonprofit hospitals was more than \$6 billion in 2002, and the value of state tax exemptions was also more than \$6 billion. *Id.* at 5, Table 1 (citing JCT as its source); and report No. 2007-10-061 of the Treasury Inspector General for Tax Administration (Mar. 29, 2007).

³¹ CBO report, *supra*, at 4. Of particular interest are the references given in fn. 5. Fn. 5 cites the 2006 GAO report on community benefits, the book chapter by Sloan FA in *HANDBOOK OF HEALTH ECONOMICS* (2000), and S/G.

The S/G article's overall conclusions

The article's conclusion stated in the S/G Abstract are biased (and therefore not objective) in indicating that the information in the article is "the most recent and comprehensive evidence", and that the article's analysis demonstrates that charges leveled at nonprofit hospitals are either wrong or incomplete in indicating that they do not reliably provide community benefits commensurate with their tax subsidies. The information in the article (including its Technical Appendix) indicates wide inconsistency among the hospital studies, and does not support such a conclusion. Overall, the article is also biased and inaccurate in implying that "quality" and cost³² differentials are taken into account in considering "community benefits."

The S/G article is also biased (and therefore not objective) in stating that the two Devereaux et al. reviews, discussed below, found "consistent" ownership-related differences, and in stating with certainty that those two studies found higher mortality rates and higher prices in for-profit hospitals. (At W289)

2. Devereaux et al. 2002 (mortality differences)

- The synthetic relative risk (RR) derived by the study is so weak as to be virtually meaningless.³³ The synthetic RR is 1.020, with 95% confidence intervals (CI) of 1.003 to 1.038. Such a synthetic relative risk is so fragile that, for example, a single added study³⁴, a small bias in one or more studies, a failure to adjust for a confounder or residual confounding, could change the outcome in the direction of advantage to for-

³² The use of "cost" as a measure is misleading. "Cost" is considered cost to the hospital of providing the service, as distinguished from "charges" or "prices" (what is billed for the services) and "payments" (what the hospital actually recovers after billing).

³³ See, e.g., Weed DL. 2000. Interpreting epidemiological evidence: How meta-analysis and causal inference methods are related. *Int'l J Epidem* 29:387-90, and references therein. ("There is an important distinction to be made between improving the precision of the relative risk estimate -- which meta-analysis offers -- and interpreting the causal relevance of the absolute value of the summarized estimate -- which it does not. [References omitted.] . . . [A] summary estimate of 2.0 -- whether it emerges from a meta-analysis or not -- remains on the borderline of what is typically called a 'weak' association."); Federal Focus. 1999. Epidemiology in hazard and risk assessment: A partial review of the "London Principles" by an expert academic panel (Berga SL, Daling JR, Friedenreich C, Greenland S, Savitz DA, Waddell WJ) (LOC Cat. No. 99-76949, ISBN 0-9654148-1-7) at 32 (comments that a particular meta-analysis with a synthetic RR of 1.29 was very weak and would be regarded by some as reassurance that there is no risk).

³⁴ E.g., Thornlow DK and Stukenborg GJ. 2006. *Med Care* 44(3):265-69 ("Hospital ownership and teaching status is not a consistent predictor of differences in rates of potentially preventable adverse events, and these characteristics explain little of the observed variation in the rates of these events across hospitals.")

profits.³⁵ In addition, any such analysis is likely to become quickly outdated due to ongoing changes in the hospital sector. A recent study found that overall U.S. hospital quality was improving between 2002 and 2004.³⁶

- As an example of a flaw that could easily have influenced the meta-analysis and produced a misleading small positive RR in favor of nonprofits, the analysis did not adequately consider or account for a major potential source of confounding: socioeconomic status (“SES”). The 2006 CBO study³⁷ found that “[o]n average, nonprofit hospitals were found to operate in areas with higher average incomes, lower poverty rates, and lower rates of uninsurance than for-profit hospitals.” (At 1.) It is widely recognized that socioeconomic status is a confounder for mortality rates³⁸, and it is almost certain that failure to take into account that confounder would have biased the results in favor of average lower mortality at nonprofit hospitals. Although Devereaux et al. recognized that it was “appropriate” to adjust for SES, the online technical appendix for the article (Table 3) shows that only five of the 13 studies examined utilized any adjustment for SES, and for some it appears to have been incomplete.³⁹ Of those five included studies, three produced non-significant results, one indicated a negative association between nonprofit status and mortality, and one indicated a positive association. The failure to base the meta-analysis only on studies with appropriate adjustment for SES represents a failure to use “sound statistical and research methods”, in violation of the Data Quality objectivity standard.

³⁵ The study Abstract states that the results “suggest” that for-profit status results in higher mortality; however, in the remainder of the study, and in the references back to this study in the 2004 study, the results are stated without this qualification.

³⁶ Williams SC et al. 2005. Quality of care in U.S. hospitals as reflected by standardized measures, 2002-2004. *NEJM* 353:255-64 (utilizing new standardized performance indicators for acute myocardial infarction, heart failure, and pneumonia).

³⁷ Fn. 28, above.

³⁸ E.g., Shishehbor MH et al. 2006. Association of socioeconomic status with functional capacity, heart rate recovery, and all-cause mortality. *JAMA* 295:784-92; Alter DA et al. 2006. Socioeconomic status and mortality after acute myocardial infarction. *Ann Intern Med* 144(2):82-93; Steenland K et al. 2004. Individual- and area-level socioeconomic status variables as predictors of mortality in a cohort of 179,383 persons. *Am J Epidemiol* 159:1047-56; Steenland K et al. 2004. All-cause and cause-specific mortality by socioeconomic status among employed persons in 27 US States, 1984-1997. *Am J Pub Health* 94(6):1037-42; Lantz RM et al. 1998. Socioeconomic factors, health behaviors, and mortality. *JAMA* 279(23):1703-08.

³⁹ SES can encompass not only education and income, but also factors such as marital status, family environment, occupation, community environment, and ethnicity. When graded on an area-wide basis, SES usually reflects income and education.

Center for Regulatory Effectiveness

The Devereaux et al. article did not include or comment on an available study that expressly considered how location might affect average differences in mortality between for-profit and nonprofit hospitals.⁴⁰ The authors of that study concluded:

We find that, on average, for-profit hospitals have higher mortality among elderly patients with heart disease, and that this difference has grown over the last decade. However, much of the difference appears to be associated with the location of for-profit hospitals. Within specific markets, for-profit ownership appears if anything to be associated with better quality care. Moreover, the small average difference in mortality between for-profit and not-for-profit hospitals masks an enormous amount of variation in mortality within each of these ownership types. Overall, these results suggest that factors other than for-profit status per se may be the main determinants of quality of care in hospitals. [Abstract]

The failure by Devereux to consider such a study and the factors it examined is a strong indication of bias/lack of objectivity (particularly in the form of lack of completeness) under the Data Quality standards.

- One of the other fundamental weaknesses of the study is that it acknowledges substantial unexplained heterogeneity⁴¹ that largely undercuts its synthetic findings, but then attempts to justify its finding of a statistically significant synthetic RR on policy grounds. The authors state:

When studies show important differences in results [as here], rigorous systematic reviewers explore the data to see if they can identify cogent explanations for the differences. How they should proceed if they fail to find an explanation for the differences remains controversial. Some argue that under these circumstances, pooling is inappropriate. Others argue that clinicians, and in this case health-policy makers, must still make decisions, and their decisions should be driven by the best available estimate of treatment effect. In the presence of unexplained heterogeneity, while inferences associated with pooled estimates are weaker, these estimates nevertheless provide the best estimate about the average effect, and thus constitute useful information for decision-makers.

⁴⁰ McClellan MB and Staiger D. 1999. Comparing hospital quality at for-profit and not-for-profit hospitals. NBER (National Bureau of Economic Research) Working Paper No. W7324, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=197072.

⁴¹ "Heterogeneity" often refers simply to differences in study outcomes (*i.e.*, relative risk (RR) calculations and statistical significance), but it also refers to differences in study design and quality -- *e.g.*, populations studied, sample size, adjustments made for confounders, case-control vs. cohort).

Center for Regulatory Effectiveness

This is an infusion of policy bias into the study, in violation of the Data Quality objectivity standard. The authors present an extremely weak synthetic RR, acknowledge that even that RR is further weakened by unexplained heterogeneity, then seek to justify it on policy grounds.⁴²

- The study is misleading, and therefore biased, in stating that the studies examined observed similar populations. By "similar populations" it appears the article is referring only to studies divided according to whether they involved adults or children. In view of the failure to adequately account for socioeconomic differentials and types of patient populations, that statement is inaccurate.

3. Devereaux et al. 2004 (charges for care)

- At the outset, it should be noted that this study contains references back to the 2002 mortality study discussed above, but rather than acknowledging the weakness of the results and stating that they only "suggest" a relationship, this study states that the 2002 study has "shown" and "demonstrated" higher mortality at for-profit hospitals.⁴³ This is a clear instance of bias.
- As with the 2002 study, the synthetic relative risk derived in this study was extremely weak: RR of 1.19 with CI of 1.07 - 1.33. The authors state that this "demonstrated" that for-profits are associated with higher payments for care and that the evidence "strongly

⁴² See Federal Focus, *supra* n. 33, at 49 (consensus on Principle II, regarding meta-analysis, stating in part: "When the evidence is conflicting, the primary goal of a meta-analysis should be identifying and estimating differences among study-specific results, rather than estimating an overall quantitative summary of the association across studies. With this goal, problems in combining evidence from studies that are dissimilar in design and quality can be minimized.") See also the discussion at 50. This Principle was an edited version of a principle regarding meta-analysis (B-6, at 57-59) drafted by another, earlier, expert panel convened by Federal Focus. The explanation of that principle stated that it incorporated a standard "that the reviewer should examine the individual studies carefully for quality and for explanations for study differences rather than combining study results simplistically to get a single point estimate." *Principles for evaluating epidemiologic data in regulatory risk assessment, developed by an expert panel at a conference in London, England, October 1995* (LOC Cat. No. 96-88988, ISBN 0-9654148-0-9). See also the consensus statement by another expert panel, stating: "In cases when heterogeneity of outcomes is particularly problematic, a single summary measure may well be inappropriate" Stroup DF et al. 2000. Meta-analysis of observational studies in epidemiology: A proposal for reporting. *JAMA* 283(15):2008-12. (The consensus statement also commented: "The application of formal meta-analytic methods to observational studies has been controversial. One reason for this has been that potential biases in the original studies, relative to the biases in RCTs [random controlled trials] make the calculation of a single summary estimate of effect of exposure potentially misleading. Similarly, the extreme diversity of study designs and populations in epidemiology makes the interpretation of simple summaries problematic, at best. In addition, methodologic issues related specifically to meta-analysis, such as publication bias, could have particular impact when combining results of observational studies." At 2011, citations omitted.)

⁴³ At 1817 and 1822 (including the first sentence of the Abstract).

Center for Regulatory Effectiveness

supports" a policy of nonprofit health care delivery. (Abstract) This also is a clear instance of bias.⁴⁴

- As with the 2002 study, the authors acknowledge the presence of substantial unexplained heterogeneity among the individual studies, which obviously further weakens any inference that might be drawn from the synthetic RR; but also, as in the previous study, this weakness is dismissed on policy grounds, and this time the explanation of the heterogeneity problem is watered-down considerably, with the authors stating that it is "common practice" to pool results with significant heterogeneity, and that the conclusion of higher payments with for-profits is "secure", with just the magnitude of the effect in question as a result of the heterogeneity. This is clearly a biased statement.
- In view of the weakness and inconclusiveness of the study findings, the extrapolation of the synthetic RR to generate nationwide expenditures of \$3.7 billion as a firm number is unwarranted, and is another instance of bias.
- One factor not considered in the study that could have strongly influenced the synthetic RR is that nonprofit hospitals have a substantial motivation to report uncompensated care, including bad debt, in order to demonstrate community benefit to maintain their tax-exempt status. The study also failed to consider the related point that for-profits might simply be more motivated and efficient in collecting charges for services.
- It is not clear that the methodology employed in the meta-analysis to generate even the weak synthetic result is valid. The comparisons were based on "case mix", which is not defined but appears to represent the case mix for the entire facility. A comparison based on payments for the same treatment (or even the same disease) at different facilities might be valid; but a comparison based on case mix does not appear valid. In addition, as reflected in Table 3, two of the studies that reported higher payments at for-profits (Van Ness and Dickey) did not adjust for case mix, whereas the two studies that reported lower payments or non-significant difference did adjust for case mix. Moreover, it is not clear how matching could be accurate in the case of psychiatric disorders (McCue et al.), since each patient would be unique (as with other disorders, but more so). And, the weighting of the individual studies was based on number of facilities, rather than number of patients or payments received, which easily could have distorted the results. All of these methodological shortcomings raise questions as to whether the study employed "sound statistical and research methods" required for "objectivity".
- At the end of the article, the authors state: "For-profit hospitals **result in** [*i.e.*, cause] both higher mortality rates and greater payments for care than do not-for-profit hospitals. The evidence **strongly** supports a policy of not-for-profit health care delivery at the

⁴⁴ At the end of the article, however, they state that their findings "suggest" higher payments at for-profits. At 1823.

Center for Regulatory Effectiveness

hospital level."⁴⁵ This concluding statement regarding both studies is perhaps the strongest evidence of bias (in violation of the objectivity standard).⁴⁶

Conclusions

Review of these three studies indicates why it is important to examine carefully the quality of studies even when they have been published in peer-reviewed journals. It is clear that journal peer-review, as illustrated here, does not suffice to establish that a study meets Data Quality standards.⁴⁷ It is also clear that such studies do not come close to resolving the ongoing issues underlying the current nonprofit hospital tax exemption.

All three studies are vulnerable to challenges based on substantial Data Quality deficiencies, particularly with regard to objectivity and the reliability of their methodology. If an IRS information dissemination were to rely on any of these three studies, it would be subject to a Data Quality challenge. We believe that IRS pre-dissemination review for compliance with Data Quality standards would confirm these deficiencies.

* * *

If you have questions concerning this information, or wish to discuss it further, please feel free to contact me at (202) 265-2383 or tozzi@thecre.com.

We will look forward to any further reports by the IRS on its plans for analyzing or acting on the overall issue of tax exemption guidance or rules for hospitals.

Sincerely,

/s/

Jim J. Tozzi
Member of the Advisory Board

Cc: Hon. J. Russell George, Treasury Inspector General for Tax Administration
Stephen T. Miller, Commissioner, IRS Tax-Exempt and Government Entities Division
Lois G. Lerner, Director, Exempt Organizations Division, IRS TE/GE

⁴⁵ At 1823 (emphasis added).

⁴⁶ It should not be overlooked that the two Devereaux et al. studies were conducted by Canadians and published in a Canadian journal during an intense Canadian debate over its national health care system, and whether it should permit entry of, and competition from, for-profit community hospitals.

⁴⁷ Even if journal peer reviewers make important criticisms of a submitted manuscript, the journal editors make the ultimate decision on whether to publish and whether to require corrections or modifications. Journals typically do not make peer review comments available to the public and the scientific community.