Healthy control subjects are poorly defined in case-control studies of irritable bowel syndrome

Shireen Ghorbani, Amir Nejad, David Law, Kathleen S. Chua, Meridythe M. Amichai, Mark Pimentel
Cedars-Sinai Medical Center, Los Angeles; Harbor UCLA Medical Center, Torrance, California, USA

Abstract
Background Case-control studies are vital for understanding the pathophysiology of gastrointestinal disease. While the definition of disease is clear, the definition of healthy control is not. This is particularly relevant for functional bowel diseases such as irritable bowel syndrome (IBS). In this study, a systematic review formed the basis for a prospective study evaluating the effectiveness of commonly used techniques for defining healthy controls in IBS.

Methods A systematic review of the literature was conducted to identify case-control studies involving functional gastrointestinal disorders. “Lack of Rome criteria”, self-description as “healthy” and the bowel disease questionnaire (BDQ) were common methods for identifying healthy controls. These 3 methods were then applied to a cohort of 53 non-patient subjects to determine their validity compared to objective outcome measures (7-day stool diary).

Results “Lack of Rome criteria” and “healthy” self-description were the most common methods for identifying healthy control subjects, but many studies failed to describe the methods used. In the prospective study, more subjects were identified as non-healthy using the BDQ than using either lack of Rome criteria (P=0.01) or “healthy” self-description (P=0.026). Furthermore, stool diaries identified several subjects with abnormal stool form and/or frequency which were not identified using lack of Rome criteria or the “healthy” question. Comparisons revealed no agreement (κ) between the different methods for defining healthy controls.

Conclusions The definitions of healthy controls in studies of functional bowel diseases such as IBS are inconsistent. Since functional symptoms are common, a strict definition of “normal” is needed in this area of research.

Keywords Irritable bowel syndrome, functional GI disorders, healthy control


Introduction

Patients with gastrointestinal (GI) diseases are often studied and compared to healthy control subjects. The ability to evaluate disease and response to treatment is heavily dependent on this case-control study design and its accepted definition of “control subject”. Data from these studies can be used to support or negate the role of certain factors in the etiology of a disease or condition [1]. Without the proper definitions, the quality of data comes into question and may lead to misinterpretation of results [2].

In gastroenterology, the line between normal and abnormal can be vague, and this is a particularly significant issue in the study of functional disorders such as irritable bowel syndrome (IBS). The diagnosis of IBS is largely subjective and based on symptoms rather than physical findings [3]. Furthermore, functional bowel symptoms are common in the general population and can vary over time [4-6]. Thus, it is important when performing case-control studies in IBS that control subjects be truly healthy in order to provide true and accurate results [7].

Although 6-20% of the population suffers from IBS [8], our vague clinical definition of IBS makes the diagnosis challenging. This lack of clarity initially framed a concept that IBS was a “diagnosis of exclusion”. Over the last two decades, the Rome inclusion criteria were developed and modified in order to help enroll IBS subjects in clinical trials [9], but
still are not a definitive diagnosis of IBS. In fact, these criteria cannot in and of themselves distinguish IBS from inflammatory bowel disease or celiac sprue [10-15]. Even more importantly, the Rome documents do not provide guidance on how to determine normal subjects for comparison to IBS.

While textbook definitions of normal bowel habits range between 3 bowel movements per week to 3 bowel movements per day [16], this definition is unidimensional since there is no determination of change in function and no accounting for stool form. Therefore, there are a variety of perceptions of normal in the clinic, in research, and by the patient themselves. Given these issues, it is vital to define “normal” in case control studies.

In this study, we performed a systematic review of the literature to determine commonly used methods for defining “normal” in studies comparing to patients with functional disease. Using the results of this review, we then performed a prospective study to determine the validity of these commonly used methods to enroll healthy control subjects.

Materials and methods

Study overview

This is a two-part study consisting of a systematic review and a prospective observational study. The systematic review was conducted first, and aimed to determine the methods used to identify healthy control subjects in functional GI case-control studies. Next, a prospective observational study was designed to test the validity of the most common methods identified by the systematic review to define healthy controls, as well as the less commonly used bowel disease questionnaire (BDQ).

Systematic review

A review of the GI literature from 1950 to April 2010 was conducted in Ovid MEDLINE and PubMed, which evaluated the methods used to define healthy controls in studies of functional GI disease. Search terms included "normal stool/bowel frequency", "normal stool/bowel form", "normal stool/bowel consistency", and "functional gastrointestinal disorders", which were cross-referenced with the terms "healthy controls" and "healthy bowels". The search was limited to studies of humans and English language only. In MEDLINE, the term "feces" was substituted for "stool" as a pre-determined subject heading. Title, abstract, and full paper reviews identified papers for final inclusion based on confirming: 1) a control group with ≥10 subjects; 2) a comparator group of confirmed disease; and 3) the paper was not a review article. After inclusion and exclusion criteria were applied, final paper selections were reviewed in detail. The studies were examined for the terminology described to define their normal populations. Furthermore, efforts to ensure normalcy of control subjects were identified and summarized.

Prospective study

Non-patient subjects between the ages of 18 and 65 were eligible to participate and asked to complete a questionnaire regarding their bowel function, habits, and symptoms followed by a stool diary. Subjects were excluded if they had a history of major abdominal surgery, known GI or liver disease, diabetes mellitus, uncontrolled thyroid disease, and taking chronic medications known to affect gut function. Subjects were recruited from August 2010 to November 2010. This study was approved by the Cedars-Sinai Institutional Review Board.

Questionnaire packet

The systematic review of the literature identified the most common methods for identifying healthy subjects for case-control studies of functional GI disorders. The first and most common method was that patients were defined as “healthy” when they did not satisfy the inclusion criteria for the disease in question (in this case, the Rome criteria for IBS). The second was that subjects simply had to self report that they were “healthy”. A third, more detailed method was the use of the BDQ, whereby “normal” was defined based on lack of reported symptoms after completing the questionnaire [17].

Based on these findings, subjects entering the prospective study were randomly assigned to one of the following questionnaires to complete as their first task (random assignment was done by random number generator for packets).

1. Rome III criteria for IBS

Published in 2006, the ROME III criteria [9] are symptom-based criteria used in the diagnosis of IBS. Subjects passed as “healthy” if they answered “no” to question 1 or if they answered “yes” to question 1 but answered “no” to two out of the 3 components of question 2.

1) Do you have recurrent abdominal pain or discomfort for ≥3 days per month in the last 3 months?
2) Do you suffer with any of the following:
    a. Improvement in pain or discomfort with defecation?
    b. Onset of pain or discomfort associated with a change in stool frequency?
    c. Onset of pain or discomfort associated with a change in stool form (appearance)?

2. “Healthy” question

Subjects were asked the simple question: “Do you consider yourself to have healthy bowel function?” If they answered yes, they were considered to be “healthy.” If they answered no, they were considered to be “not healthy” and were not eligible to be a healthy control.

3. BDQ

The original BDQ [17] was developed to distinguish patients with functional GI disorders from those with organic GI conditions, by addressing 46 GI symptoms among other symptoms and health problems. Nineteen of the 46 symptoms correlated with IBS and bowel habits, and these were selected and compiled to create a modified...
version of the BDQ used in this study. Subjects were deemed “healthy” if they scored <4 positive GI symptoms as being of “mild” severity at worst. This distinction was based on the application of the BDQ by Dieteren et al, who used these criteria to define healthy controls in their study [18].

**Validation of primary questionnaire**

After subjects completed their randomly assigned primary questionnaire, a second questionnaire in a sealed envelope was opened and completed by all subjects. The second questionnaire had 3 sections. In the first section, subjects were assessed using the 2 methods by which they were not initially assessed. For example, subjects who initially completed the Rome III criteria (the “Rome” group) went on to answer the “Healthy” question and the BDQ as their secondary outcome measures. Likewise, subjects who completed the “Healthy” question first (the “Healthy” group) went on to answer “Rome” and “BDQ” as their secondary outcome measures.

The final portion of the second questionnaire was a stool diary. Subjects were asked to complete a 7-day stool diary which evaluated bowel consistency based on the Bristol Stool Scale, as well as frequency, ease of passage, and bowel evacuation.

**Data analysis**

The initial questionnaire was first assessed for its ability to eliminate subjects as “not healthy” from a functional bowel standpoint. This determination was then compared to subjects’ responses to the remaining question groups. Finally, the bowel symptom questionnaire and stool diary were used to determine whether subjects were truly normal or did in fact suffer from functional GI symptoms.

**Statistical analysis**

The Fisher’s exact test was used to compare the proportions of abnormal subjects identified by each of the three methods for defining a healthy control. To compare agreement between methods, Cohen’s κ was employed.

**Results**

**Systematic review**

A total of 340 abstracts were identified using the initial search terms. After applying the inclusion and exclusion criteria, 43 papers met the criteria for the systematic review comparing controls to functional GI disease [18-60].

It was discovered that the most common method of defining a healthy control was determining that an individual did not meet the criteria for the disease to be studied (Table 1). For example, in a study comparing healthy control subjects to IBS patients, an individual would be deemed to be a control if he/she did not meet the Rome criteria for IBS. The second most common method of defining a healthy control was to simply ask an individual if they considered their bowel function as “healthy”, without further evaluation being mentioned. In addition, there was a number of studies that did not describe the methods used to identify healthy controls at all (n=7). Thus in over 80% of studies, the definition of normal was not based on detailed criteria or responses.

In approximately 15% of studies, the definition of a healthy control was based on detailed stool form and frequency identified using the BDQ. However, even these studies exhibited considerable heterogeneity as to what that definition of “normal” was. Based on this review, exclusion of Rome criteria, “healthy”, and the use of the BDQ to define normal were the methods most commonly used to define healthy controls.

**Prospective study**

**Study subjects** Sixty subjects were recruited for the study, of whom 53 met the inclusion and exclusion criteria and were enrolled in the study. Of these, 14 were randomized to the “Rome” group, 19 to the self-defined “healthy” group, and 20 to the “BDQ” group. Of the 53 subjects, 21 were female, 31 were male, and 1 did not specify. The mean age of study participants was 31.6±10 years old.

**Primary questionnaire** Although the subjects were recruited as non-patients, 7 of 20 subjects in the “BDQ” group were deemed to have IBS based on the modified BDQ (Fig. 1). This would have rendered them ineligible as healthy controls for case-control studies. This was significantly greater than the “healthy” group, where only 1 subject identified their bowel function as not healthy (P=0.026), and the “Rome” group, where no subjects were identified as IBS (P=0.01 compared to “BDQ”). Thus, the modified BDQ detected functional symptoms in non-patient subjects with a greater frequency than the “healthy” question or the Rome Criteria.

**Comparison of the three methods for the identification of healthy controls** Comparing the 3 above-described methods demonstrated significant inconsistency in identifying healthy controls (Table 2). Of the 18 subjects defined as normal

---

**Table 1** Methods used to define normal controls identified by the systematic review

<table>
<thead>
<tr>
<th>Definition of control</th>
<th>Functional disease (N=43) N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not meet criteria of comparator disease¹</td>
<td>17 (39)</td>
</tr>
<tr>
<td>“Healthy”</td>
<td>6 (14)</td>
</tr>
<tr>
<td>Not defined</td>
<td>7 (16)</td>
</tr>
<tr>
<td>Combination</td>
<td>7 (16)</td>
</tr>
<tr>
<td>Utilized Bristol Stool Score</td>
<td>4 (9.1)</td>
</tr>
<tr>
<td>Bowel range</td>
<td>2 (4.5)</td>
</tr>
</tbody>
</table>

¹In most cases this meant subjects did not meet the Rome Criteria for irritable bowel syndrome.
based on the “healthy” question, 7 (39%) would be defined as not normal and excluded using the modified BDQ. Likewise, of the 14 subjects defined as normal using the Rome criteria, 4 (29%) would be defined as not normal and excluded using the modified BDQ. In contrast, the Rome criteria and the “healthy” question never altered the assignment of “BDQ” subjects.

On the basis of abdominal pain alone, 3 of the 20 subjects in the BDQ group would be categorized as IBS per Rome criteria; all 3 were already defined as abnormal by the modified BDQ. In contrast, the Rome criteria and the “healthy” question each only identified 1 subject as abnormal. Agreement between all three methods for identifying healthy controls and the Bristol Stool Scores was also examined, and no agreement was found between any of the methods used (Table 3).

**Table 2 Subject characterization on primary questionnaire vs. other methods**

<table>
<thead>
<tr>
<th>Primary outcome</th>
<th>BDQ</th>
<th>Healthy question</th>
<th>Rome criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Functional</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Healthy</td>
<td>18</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Not healthy</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

**Figure 1** Proportion of non-patients who would not meet criteria for healthy control based on their assigned primary questionnaire

<table>
<thead>
<tr>
<th>BDQ, bowel disease questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of subjects not deemed normal</td>
</tr>
<tr>
<td>0%</td>
</tr>
</tbody>
</table>

**Discussion**

This is the first study to systematically review the commonly used approaches to identify healthy controls in IBS studies and to evaluate their validity. As evidenced by the data presented, there is no standardized definition of a “healthy control” regarding bowel function. Many of these studies use failure to meet the Rome criteria, or self-definition as “healthy” as qualifying measures for their controls. In our study, the less-frequently used BDQ was most successful at distinguishing subjects with normal vs. abnormal bowel function (as determined using 7-day stool diaries), but there was no consistency between the techniques used.

Since IBS and other functional GI disorders have largely remained diagnoses of exclusion, it is necessary to continue research studies aimed at identifying the underlying etiologies or predisposing factors. Case-control studies are very important for this, and provide a foundation on which randomized controlled trials can be built. Recent studies of IBS patients versus controls have examined the role of GI transit and hormones in IBS symptoms [19,22,36,41], anorectal function...
Table 3 Agreement (k) between methods for defining healthy controls

<table>
<thead>
<tr>
<th></th>
<th>BDQ question</th>
<th>Healthy question</th>
<th>Rome criteria</th>
<th>Bristol score</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDQ</td>
<td>0.07</td>
<td>0.22</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>Healthy question</td>
<td></td>
<td>0.22</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Rome criteria</td>
<td></td>
<td></td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Bristol score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BDQ: bowel disease questionnaire

in IBS [30,33], quality of life issues [20,31], and presence of methanogenic flora on breath testing [42,47]. These studies and similar ones have led to advancements in our understanding of the etiologies underlying IBS and therapeutic and behavioral measures that can be taken for symptom improvement. Essential to these case-control studies is an appropriately identified group of healthy controls for comparison, defined by what is accepted as "normal" for that particular entity. Our systematic review illustrates that, in the case of IBS, this is usually done by choosing individuals who do not meet the Rome criteria or who self-report as "healthy". We must then look at what has commonly been regarded as healthy bowel habits.

The generally accepted range of "healthy" is having between 3 bowel movements per day to 3 bowel movements per week. This designation dates back to 1965, in an observational study of 1055 adults by Connell et al [61], who reported that 99% of individuals fell into the aforementioned range, and any frequency which deviated from this was considered abnormal. Notably, this definition does not account for other qualities of stool, such as form, consistency, and associated GI symptoms, and leaves a gaping hole in our already substandard understanding of healthy bowel habits. Would an individual be considered normal if they had 3 bloody bowel movements per day, or 3 bowel movements per week associated with straining and abdominal cramps? Over the years there has been some focus on these additional factors, however, no standard definition has emerged [62,63].

Similarly, it is known that functional GI symptoms are common in the general population and wax and wane over time, whether in IBS patients or not [4-6]. In fact, so reliable is this characteristic that it has been used to establish a new clinical method to discriminate between diarrhea-predominant IBS and non-IBS patients [64]. The high prevalence of these symptoms complicates the selection of truly healthy controls in IBS studies, as many individuals who self-report as healthy may actually have an underlying disorder. Indeed, an individual may perceive having numerous loose stools on a daily basis as normal. Such is the intrinsic flaw in recruiting control patients on the basis of self-reported healthy bowel habits or symptom-based criteria. In this study, we note that a number of our non-patient subjects had IBS-like features with abnormal bowel function and should not be enrolled as "healthy controls" in studies of functional bowel disorders, but would have been included using most of the commonly used methods for identifying control subjects.

In our study, comparing the three most commonly used methods to define healthy controls in IBS case-control studies (failure to meet Rome criteria, self-definition as healthy and BDQ) revealed marked inconsistency between these methods as well as a general mismatch between the results and an individual's actual and perceived bowel habits. Using self-reports of healthy bowel function versus the modified BDQ made it far more likely to miss a diagnosis of IBS. Applying that concept to the whole of IBS literature suggests that many control subjects may actually have IBS or other underlying GI conditions, undermining the very principle of a case-control study.

In addition to having poorly defined healthy controls, our review of these case control studies revealed other systemic flaws. IBS rates are predominantly determined by health-care seeking behaviors, thus making age- and gender-matching vital to appropriate interpretation of studies. Many studies however do not appear to adequately match healthy controls to IBS subjects. In a study by Bratten et al, investigators not only provide little or no definition of healthy controls, but also reveal the mean age of controls (<20 years old) to be more than a decade younger than enrolled IBS subjects [42]. This control group is not even likely to readily seek health care as they are not yet gainfully employed, and likely do not seek routine health care (e.g. well-woman visits which begin at a later age) in which review of symptoms might identify the common condition of IBS.

There are pitfalls in this study. The sample population consisted mostly of hospital employees, and may not be an accurate representation of the general population in terms of bowel habits, daily activities, and stress levels. Study subjects may have differing interpretations of the Bristol Stool Form Scale, which may cause a false normal/abnormal result. Likewise, GI symptoms in one individual may not be thought of as bothersome in another individual and therefore, may not be reported as such.

The pathophysiology of IBS is still poorly understood by the gastroenterology community, despite years of research aimed at identifying the underlying etiologies. In addition to continuing research on potential mechanisms leading to IBS, it is essential to evaluate the role that healthy control subjects play in IBS studies. In the absence of a uniform definition of normal bowel function, the interpretation of the GI literature in case-control studies is suspect. IBS research would benefit from a validated tool, perhaps one which focused on abdominal pain (quality, location, severity) and bowel habits (form, frequency, straining), to standardize healthy control selection. The progress in studying pathophysiologic processes in IBS will mandate a uniform definition of healthy controls. Based on this study, there is no endorsed accurate method for identifying a healthy control for case-control studies.

Acknowledgments

We would like to thank the Beatrice and Samuel A. Seaver Foundation for their ongoing support of our research in the GI Motility Program at Cedars-Sinai Medical Center.

Annals of Gastroenterology 27
Summary Box

What is already known:

- Case-control studies are vital for understanding the pathophysiology of gastrointestinal disease.
- While the definition of the disease under study is usually clear, the definition of a ‘healthy control’ is not.
- This is particularly relevant for functional bowel diseases such as irritable bowel syndrome.

What the new findings are:

- Based on a review of the literature, “lack of Rome criteria”, self-description as “healthy” and the bowel disease questionnaire (BDQ) were the 3 most common methods for identifying healthy controls in published studies.
- In our prospective study, more subjects were identified as non-healthy using the BDQ than using either “lack of Rome criteria” or self-description as “healthy”.
- Stool diaries identified several subjects with abnormal stool form and/or frequency which were not identified using “lack of Rome criteria” or self-description as “healthy”.
- Given these inconsistencies in the definitions of healthy controls, a strict definition of “normal” is needed in this area of research.

References

Defining healthy controls compared to IBS


