Background:
In advance of a new medication being introduced to market, it is important to understand current patterns of care and associated clinical outcomes, in order to document current treatment gaps and the new therapy will address.
Retrospective chart review is a useful methodology for efficiently addressing these types of questions. A limitation of chart review studies is that the process of extracting comprehensive data from charts can impact sample size feasibility.
In contrast to the hypothesis-testing framework, for a burden of illness study, the role of sample size is to improve precision around estimates of descriptive outcomes, and ensure sufficient representation is achieved amongst subgroups of clinical interest.

Determining how many patient charts to include in such a study is not straightforward: common sample size calculations developed for a hypothesis-testing framework have limited applicability in this context. The objective of these analyses was to develop and present rigorous approaches for sample size calculation for such studies.

Methods:

• Relevant sample size calculations are provided below for common burden of illness objectives

1. Probability of observing a treatment, and expected precision
• Characterizing the distribution of treatment patterns and precision around estimates
• Identifying all medications used in practice for an indication, including relatively infrequently used therapies (e.g. including infrequently and off-label used therapies)

2. Characterizing precision around estimates of direct medical costs, for an overall sample and key subgroups

• An illustrative case study is presented of a retrospective chart review of advanced melanoma, and the precision obtained for a sample size of 655 patients across three countries, including ability to consider patient subgroups and compare costs by country.

Results:

• For objectives around characterizing treatment patterns and distributions, the relationship between outcomes of interest and sample size are displayed in Table 1
  - For sample sizes of ≥100, confidence interval precision for treatment distribution is consistently below ±10.
  - For sample sizes of ≥100 patients and treatments given to ≥10% of the population, there is statistically a 100% chance of observing the treatment at least once.

Table 1: Relationship between sample sizes and the expected number of cases to be observed, the probability of observing a treatment in practice, and expected precision, for sample size and true probability of treatment assignment.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Expected Number of Individuals Receiving Treatment</th>
<th>Probability of Observing Treatment at Least Once</th>
<th>Precision of Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=10</td>
<td>1.0 (0.99) ±0.1</td>
<td>0.99 ±0.09</td>
<td>1%</td>
</tr>
<tr>
<td>n=50</td>
<td>1.0 (1.00 ±0.05</td>
<td>1.00 ±0.05</td>
<td>5%</td>
</tr>
<tr>
<td>n=100</td>
<td>1.0 (1.00 ±0.03</td>
<td>1.00 ±0.03</td>
<td>10%</td>
</tr>
</tbody>
</table>

For sample sizes of ≥100 patients and treatments given to ≥10% of the population, the expected confidence interval precision for treatment distribution is consistently below ±10.

• For objectives around medical cost estimation, estimated precision by sample size is straightforward to calculate if an estimate of the standard deviation is available.
  - If no estimates of the standard deviation are available, an alternative option is to consider the ratio of standard deviation to mean costs.
    - In a real-world example of observed costs for late stage melanoma patients in the UK, Italy, and France, the ratio of standard deviation to mean ranged from 0.3-4.5 across countries and outcomes, with a median of 0.7 and mean of 1.6.1

Table 2 lists required sample sizes for a range of standard deviation-to-mean ratios and desired width of confidence intervals (expressed as % of mean costs)

- Across a range of standard deviation-to-mean ratios from 0.2 to 4.5 (as observed in the real-world example), to achieve a 95% confidence interval precise to within 1% of mean costs would require more than 1,000 charts.
- Assuming a standard deviation-to-mean ratio of 1.0 (mean observed ratio in the real-world example), to achieve a 95% confidence interval precise to within 15% of mean costs would require more than 171 charts.

Note that Table 2 can also be applied to subgroups, e.g. if a subgroup that forms 50% of the population is of interest, the required sample size in Table 2 would be doubled to identify the overall required sample size to achieve the desired precision for the subgroup of interest

Table 2: Sample size required to achieve desired precision (% of mean costs) based on 95% confidence interval width of mean costs for a range of ratios of standard deviation to mean costs (y)

<table>
<thead>
<tr>
<th>Standard Deviation-to-Mean Ratio</th>
<th>Sample Size</th>
<th>95% CI Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>500</td>
<td>±10%</td>
</tr>
<tr>
<td>0.3</td>
<td>3,457</td>
<td>±1%</td>
</tr>
<tr>
<td>0.5</td>
<td>9,604</td>
<td>±1%</td>
</tr>
<tr>
<td>1.0</td>
<td>38,416</td>
<td>±15%</td>
</tr>
<tr>
<td>2.0</td>
<td>153,664</td>
<td>±30%</td>
</tr>
<tr>
<td>3.0</td>
<td>345,744</td>
<td>±45%</td>
</tr>
<tr>
<td>4.5</td>
<td>777,926</td>
<td>±60%</td>
</tr>
</tbody>
</table>

Conclusions:

• For objectives around treatment patterns and distributions, sample sizes of 100 patients and greater are likely sufficient, although larger samples may be required to characterize cost outcomes and/or examine subgroups.
• For objectives around cost estimation, greater sample sizes are required, particularly if relatively precise estimation is desired, and/or if subgroups are of interest.
• If 200-400 patient charts is the maximum feasible sample size, as is often the case in practice, it can be expected that cost estimates 95% confidence intervals will be precise to within 5-15% of the mean, depending on the ratio of standard deviation to mean costs.

This methodological study addresses an important knowledge gap, as sample sizes are frequently determined using ad-hoc approaches and/or based only on feasibility considerations.

• Formulas can be used in two distinct ways:
  - If study resources are flexible and desired precision is known, formulas can be used to guide sample size selection
  - If study resources or available sample size is fixed, formulas can be used to generate anticipated values of precision

The approach presented here is methodologically rigorous and designed for practical application in real-world retrospective chart review studies.

References: