Relevant BioStatistics Literature

For the problem of **"Doubling sample size based on a p-value,"** you need to look at literature regarding **Adaptive Designs**. If not done correctly, you inflate the "Type I Error" (false positives).

- **Key Paper:** *C. Mehta and S.J. Pocock (2011). "Adaptive increase in sample size when interim results are promising: A practical guide with examples."* Statistics in Medicine.
- **The "Promising Zone" Concept:** This paper defines exactly when you are allowed to increase sample size without "cheating" the p-value. It is the gold standard for this specific maneuver.

For the problem of the **"Internal Clinical Lead seeing unblinded data,"** the literature focuses on **Data Monitoring Committees (DMC)**.

- **Key Document:** *FDA Guidance for Industry: Establishment and Operation of Clinical Trial Data Monitoring Committees.*
- **Key Paper:** *Ellenberg, S. S., Fleming, T. R., & DeMets, D. L. (2019). "Data Monitoring Committees in Clinical Trials: A Practical Perspective."*
- **Why it matters:** It provides the "empirical evidence" that when sponsors see unblinded data early, they make biased decisions that lead to trial failure or regulatory rejection.

For the problem of **"Excluding patients on diets"** (Selection Bias), the landmark paper is the official ICH E9(R1) addendum. It changed the industry from "just use ITT" to a more sophisticated "Estimand" framework.

- **Key Source:** *ICH E9(R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials.* (EMA/CHMP/ICH/436221/2017).
- **Why it matters:** It explains why "Per-Protocol" analysis (excluding people who didn't follow the rules) is dangerous and how to use "Intercurrent Events" to stay scientifically honest.

Other Common Issues:

1. Multiplicity & "P-Hacking" (The False Discovery Guardians)

- **Benjamini, Y., & Hochberg, Y. (1995).** *"Controlling the false discovery rate: a practical and powerful approach to multiple testing."* **JRSSB** (Journal of the Royal Statistical Society: Series B).
    - **The Logic:** This is the most famous paper on "False Discovery Rate" (FDR). It proves that the old "Bonferroni" method is often too strict, but doing nothing is fatal. It provides the math for when you have dozens of endpoints.
- **Westfall, P. H., & Young, S. S. (1993).** *"Resampling-Based Multiple Testing."* **JASA** (Journal of the American Statistical Association).
    - **The Logic:** Introduces "Permutation Tests" to solve multiplicity. It catches companies that try to find a "significant" p-value by re-shuffling data until something sticks.

2. Causal Inference & Missing Data (The "Gold Standard")

- **Rubin, D. B. (1976).** *"Inference and missing data."* **Biometrika**.
    - **The Logic:** The paper that defined **MAR** (Missing At Random) vs. **MCAR** (Missing Completely At Random). If a protocol doesn't define these, the FDA assumes the data is biased.
- **Robins, J. M., et al. (2000).** *"Marginal Structural Models and Causal Inference in Epidemiology."* **Epidemiology / Biometrika.**
    - **The Logic:** The foundation of "G-estimation." It is used to fix trials where patients switch treatments halfway through (Common in cancer trials).

3. Power & Sample Size Logic (Precision vs. Luck)

- **O'Hagan, A., et al. (2005).** *"Assurance in clinical trial design."* **Pharmaceutical Statistics / Biometrika.**
    - **The Logic:** Introduces the concept of **Assurance**—the probability that a trial will actually succeed, which is more rigorous than simple "Power." It catches trials that are "technically powered" but statistically "fragile."

4. Representativeness & Generalizability

- **Pearl, J. (2011).** *"Transportability of Causal and Statistical Relations."* **JASA.**
    - **The Logic:** This is the math behind "External Validity." It proves why a trial run in one country might be scientifically invalid for a population in another country due to "Selection Bias."