The Complexion of Risk and Error in Patient Selection for CNS Clinical Trials: Detailed Findings from a 16,000-Patient Eligibility Review Database

KR Nations, CK Reinhold, K Miloslavich

INC Research, Raleigh, North Carolina

Methodological Question

What is the nature and extent of eligibility decision errors in CNS clinical trials; and how does central, collaborative review of subject eligibility protect against unwanted heterogeneity and error?

Introduction

A growing body of evidence suggests that high CNS trial failure rates are largely attributable to inappropriate patient selection (e.g., Sacks et al, 2014). CNS protocols are universally complicated, with the number of eligibility criteria ranging up to 70 in many trials. Investigators, while committed to data quality and protocol compliance, interpret protocol criteria differently and have varying levels of tolerance for risk. The subjectivity of every diagnosis in the CNS field, as well as the high number of judgment-driven decisions required to evaluate clinical assessments, naturally translates to compromised internal validity and power risks to clinical trials in the absence of careful oversight of incoming subjects (Jiang et al, 2010). Fortunately, centralized review is both operationally feasible and acceptable to Investigators, when conducted collaboratively leaving final eligibility decisions in the hands of the treating physicians.

Methods

Eligibility Review is a process conducted at INC Research by a centralized, global team of physicians and doctoral-level clinical scientists, only after Investigators have determined that they consider the subjects qualified to enter the trial (ie, subjects screen failed by Investigators are not submitted for review). Key medical and neuropsychiatric screening data are collected from sites and directly from vendor portals by project management personnel, then compiled and reviewed in a team discussion format. Resulting eligibility concerns and questions are thereafter discussed with Investigators. In the majority of cases, Investigators are able to provide additional clinical history that supports subject eligibility; however, in those cases where Investigators agree the subject is ineligible, the site proceeds to screen fail the subject. The full process and dialogue takes place within the screening period so that no ineligible patient is randomized inappropriately.

Results

To date, a total of 16,414 subjects have been reviewed for eligibility by a central medical/clinical team, covering 38 trials over a five-year period, 20 in psychiatric indications, 10 in neurology, and 8 in analgesia. Although the submitting Investigators considered all subjects to be eligible to randomize, the review team identified 1,468 (8.9%) who did not meet eligibility criteria and were ultimately considered ineligible to enter the trials. All subjects considered ineligible after review were screen failed following collaborative discussions with Investigators.

Figure 1. Rate of Findings from Centralized Eligibility Review (N Reviewed = 16,414 Subjects)

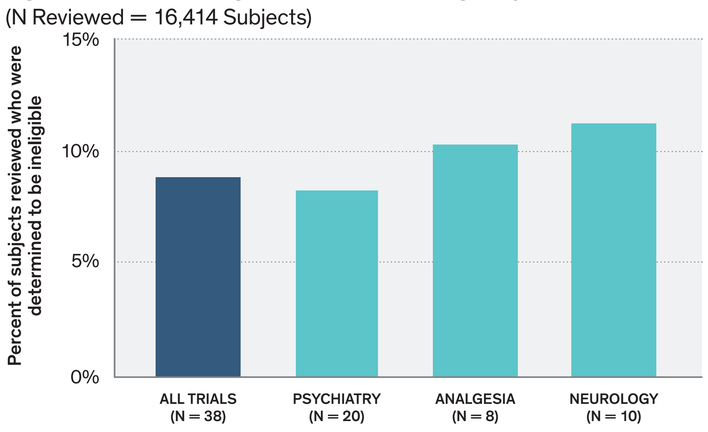

Figure shows the percent of subjects centrally reviewed who were determined to be ineligible, overall and by therapeutic area.

An exploration into to the relationship between study size and rate of findings showed that neither number of sites recruiting in the trial nor total number of subjects reviewed in the trial significantly predicted the rate of eligibility review findings (Number of sites: $R^2 = 0.01$, $p = 0.25$; Number of subjects: $R^2 = -0.03$, $p = 0.99$). .

Rate of Findings by Therapeutic Area and Indication

All analyses were undertaken in exploratory fashion without a priori hypotheses. Proportions/percentages were transformed for parametric statistical testing.

Rate of findings did not differ significantly between therapeutic areas: 8.4% of subjects reviewed were considered ineligible in psychiatry trials, 11.4% in neurology trials, and 10.4% in analgesia trials [$F(2,35) = 3.13$, $p=0.06$]. Further examination demonstrated that the rate of findings was highly dependent on individual trial, where this metric ranged from 3.7% to 25.0% in psychiatry trials, 0% to 18.1% in neurology trials, and 6.0% to 24.0% in analgesia trials. Although the rate of findings differed significantly between specific indications [$F(8,29) = 2.43$, $p=0.04$], the rate ranged widely between studies even within the same indication.

Table 1. Rate and Range of Findings: Overall, by Therapeutic Area, and by Trial Indication

	Overall rate of findings (%)	Range of findings in all trials (%)
All Trials	8.9%	0% - 25.0%
Psychiatry	8.4%	3.7% - 25.0%
Analgesia	10.4%	6.0% - 24.2%
Neurology	11.4%	0.0% - 18.1%
Major depressive disorder	9.3%	5.3% - 17.1%
Schizophrenia	5.0%	3.7% - 25.0%
Attention deficit hyperactivity disorder	7.5%	5.2% - 23.2%
Bipolar disorder	8.7%	n/a
Fibromyalgia	12.8%	11.0% - 14.2%
Neuropathic pain	21.6%	17.4% - 24.2%
Non-neuropathic pain	7.7%	6.0% - 11.9%
Alzheimer's disease	16.0%	9.6% - 18.1%
Pooled other neurology	5.5%	0% - 10.5%

Rate of Findings by Country

The overall rate of eligibility review findings differed significantly depending on country [X^2 (22, N = 16,301) = 69.94, $p < 0.001$]. Countries with more than 100 subjects reviewed tended to converge around the overall global rate (i.e., 8.9%), while countries with fewer than 100 reviews were more variable to both extremes. Nevertheless, country level differences continued to be significant even after those with <100 reviews were excluded from the analysis [X^2 (9, N = 15,584) = 35.13, $p < 0.05$].

Figure 2. Country-Level Rate of Findings from Centralized Eligibility Review (N Reviewed = 16,301 Subjects)

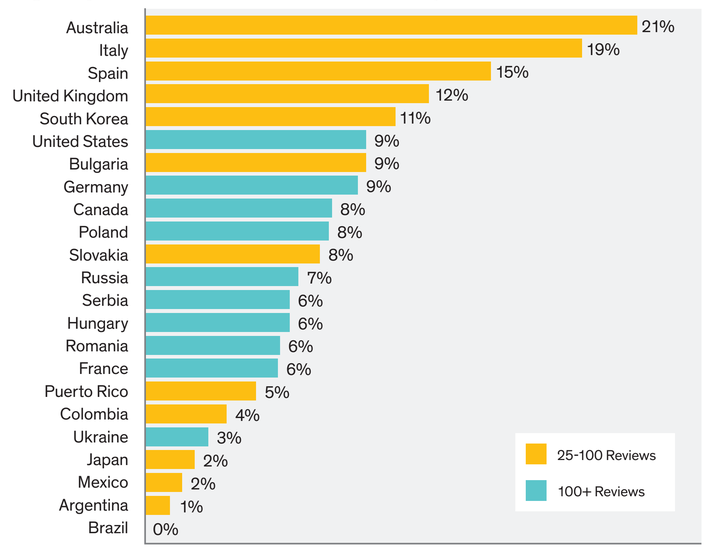

Figure shows the percent of subjects centrally reviewed who were determined to be ineligible, by country. Twelve countries with fewer than 25 subjects reviewed were excluded in support of data stability and interpretability.

Rate of Findings by Type of Site

The rate of eligibility review findings did not differ significantly between academic research centers and private research centers [X^2 (1, N = 12,919) = 2.78, $p = 0.10$].

Table 2. Rate of Findings by Type of Research Center

	Private Center	Academic Center	p Value
Sites	1,463	130	
N Reviewed	12,362	557	
Rate of Eligibility Findings	9.4%	11.5%	p = 0.10

Table shows the percent of subjects centrally reviewed who were determined to be ineligible, by type of Site. Ex-US data were excluded given the few number of academic centers utilized outside of the United States.

Clinical Category of Findings as a Result of Centralized Eligibility Review

The distribution of central eligibility review findings was broad. To date, of the subjects considered ineligible (some for more than one reason), 34.1% were ineligible for reasons related to laboratory or ECG findings, 37.3% for medical history, 27.3% for treatment history or prohibited medications, 16.5% for psychiatric history or psychosocial reasons, and 8.4% due to findings related to primary diagnostic validity.

Figure 3. Distribution of Centralized Eligibility Review Findings by Clinical Category (N Reviewed = 16,414; N Considered Ineligible = 1,468)

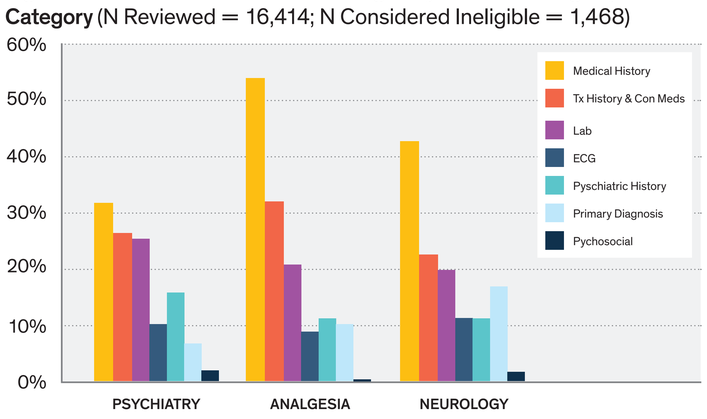

Figure shows the distribution of findings by clinical category, i.e., the reason that subjects were considered ineligible. Subjects may be ineligible for more than one reason.

Case Studies in Pursuit of Answers to Common Questions

Are sites more likely to miss on simple/explicit entry criteria or on more complex, medically nuanced protocol criteria?

A review of all eligibility findings was undertaken for a three-trial global analgesia program to better understand whether subjects submitted for centralized review were considered ineligible due to reasons that were simple protocol criteria misses (such as exceeding an explicit QTc threshold on ECG), or for reasons that were more complex, requiring consideration of multiple medical factors, further investigation of a possible exclusionary neuropsychiatric history, or safety risk (see table 3 for examples of both types). Overall, reasons for ineligibility were more often categorized as complex (73.6%) than simple/explicit (26.4%); however, simple/explicit findings were less common proportionally in the US than in ex-US countries.

Of note, misses of simple/explicit protocol criteria are those issues that would likely be detected by monitors once on site; misses of complex protocol criteria would more likely be detected by medical monitors when reviewing aggregate study data long into the trial, discovered at the time of study report or NDA preparation, or questioned by regulatory agency auditors. All findings would be considered protocol violations; however, under the process and data presented here, ineligible subjects were never randomized into the trials and therefore did not cause protocol violations.

Figure 4. Case Study Distribution of Eligibility Review Findings by Type of Protocol Criterion (N Reviewed = 1,113; N Considered Ineligible = 140)

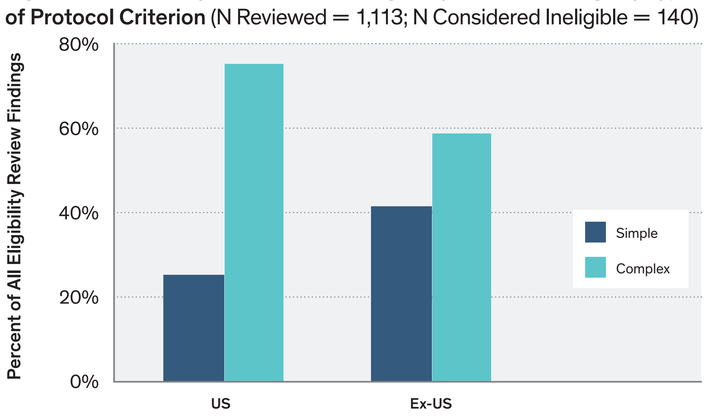

Figure shows the proportion of findings, by region, that were simple misses of explicit protocol criteria vs more complex/medically nuanced issues

Table 3. Representative Sample of Findings from Trials in Different Indications

Indication	Subject Finding	Protocol Relevance	Type
Depression	Subject with notably elevated blood pressure, labs suggestive of renal function impairment, and significant cardiac history	Exclusion criterion: unstable/ uncontrolled medical condition	Complex
Depression	Subject screening for trial while applying for disability at the same time	Exclusion criterion: issue/ motivation that could impact efficacy outcome	Complex
ADHD	Too few symptoms positively endorsed to meet DSM criteria for ADHD	Inclusion criterion: current, primary diagnosis of ADHD	Simple
Schizophrenia	Creatine kinase 6× the upper normal limit, multiple ECG abnormalities, history of hypertension	Exclusion criterion: unstable/ uncontrolled medical condition	Complex
Bipolar Disorder	Evidence of alcohol abuse detected; later confirmed severe abuse and impairment after PI further discussed with family	Exclusion criterion: recent alcohol or substance abuse or dependence	Complex
Migraine	Did not meet the minimum number of headache days per month required	Inclusion criterion: minimum headache frequency	Simple
Fibromyalgia	Current medication raised question of bipolar disorder history; past diagnosis confirmed when PI further examined medical records	Exclusion criterion: lifetime diagnosis of bipolar disorder	Complex
Alzheimer Disease	History of HTN, COPD, CAD, arteriosclerotic heart disease; clinical course and presentation of cognitive impairment more consistent with vascular dementia	Exclusion criterion: vascular or mixed dementia	Complex
Narcolepsy	Insufficient stabilization time on concomitant stimulant medication	Exclusion criterion: medication requiring pre-defined period of stabilization	Simple

Do sites improve over time with continuous feedback?

The rate of eligibility review findings was examined in two different multi-trial programs to determine whether sites show improvement throughout the course of the enrollment period. In a global psychiatry program that included approximately 150 sites and, on average, 18 subjects per site, time since the start of the enrollment period was a significant predictor of the rate of findings, where the findings rate decreased from 15.1% in the first month of enrollment to 6.2% at the end of the trial ($R^2 = 0.47$, $p < 0.01$). A global analgesia program (approximately 325 sites, an average 3 subjects per site reviewed) showed a similar trajectory of improvement over time, but in this case time was not a significant predictor of the rate of findings ($R^2 = 0.26$, $p = 0.16$). In addition, we examined the rate of eligibility review findings for sites who participated in three consecutive trials within a single psychiatry program testing the same compound. No consistent pattern emerged with regard to performance over time.

Figure 5. Within-Trial Improvement Case Study: Rate of Findings Following Centralized Eligibility Review by Month of Enrollment – A Phase III Psychiatry Program (N Reviewed ~ 2700; N Sites ~150)

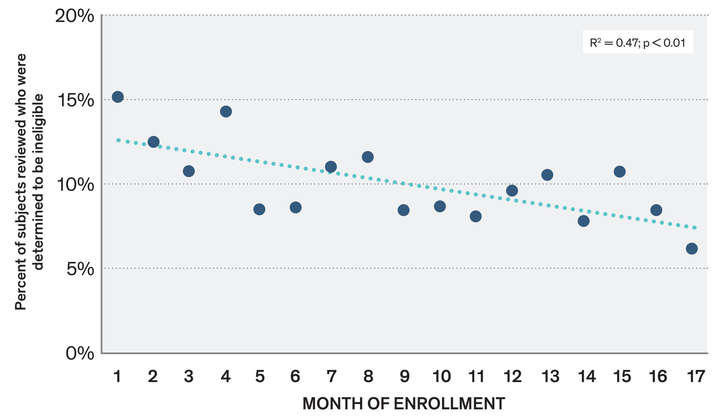

Figure shows the rate of findings, all recruiting sites combined, over the period of enrollment.

Figure 6. Within-Trial Improvement Case Study: Rate of Findings Following Centralized Eligibility Review, by Month of Enrollment – A Phase III Analgesia Program (N Reviewed ~ 1100; N Sites ~ 325)

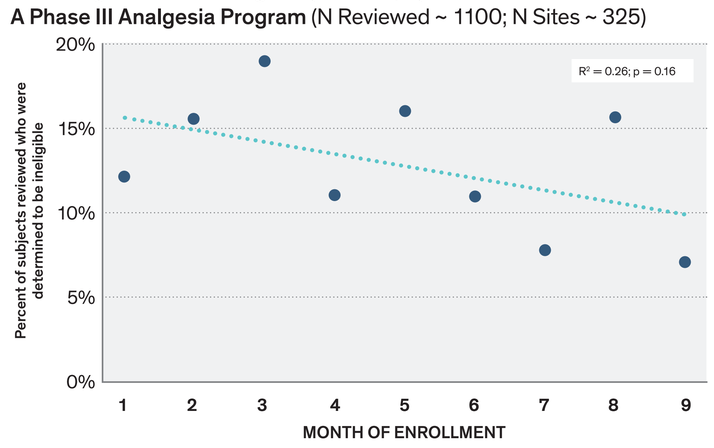

Figure shows the rate of findings, all recruiting sites combined, over the period of enrollment.

Figure 7. Between-Trial Improvement Case Study: Rate of Findings by Study for Site Investigators that Participated in Three Consecutive Trials within the Same Program

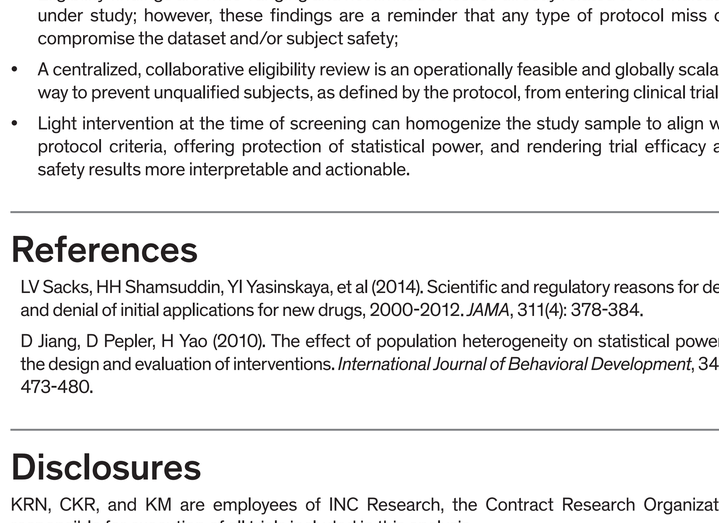

Figure shows the percent of subjects centrally reviewed who were determined to be ineligible, by study, for investigators who participated in similar, consecutive trials within the same program and submitted at least 10 subjects per study for review. Patient population was equivalent in all trials; eligibility criteria were similar.

Conclusions

The current analysis provides empirical support for the following:

- CNS trials are complex and highly susceptible to penetration by unqualified subjects;
- Experience does not protect sites against error. Academic sites are as likely to miss protocol criteria as private sites; site performance does not necessarily improve with experience in multiple similar trials. Country-level performance is highly variable and not clearly reflective of enrollment volume or experience;
- All CNS therapeutic areas and indications included in this analysis show substantial risk, albeit highly variable risk depending on individual trial;
- Eligibility findings are wide-ranging and not limited to issues directly related to the indication under study; however, these findings are a reminder that any type of protocol miss can compromise the dataset and/or subject safety;
- A centralized, collaborative eligibility review is an operationally feasible and globally scalable way to prevent unqualified subjects, as defined by the protocol, from entering clinical trials;
- Light intervention at the time of screening can homogenize the study sample to align with protocol criteria, offering protection of statistical power, and rendering trial efficacy and safety results more interpretable and actionable.

References

LV Sacks, HH Shamsuddin, YI Yasinskaya, et al (2014). Scientific and regulatory reasons for delay and denial of initial applications for new drugs, 2000-2012. *JAMA*, 311(4): 378-384.

D Jiang, D Pepler, H Yao (2010). The effect of population heterogeneity on statistical power in the design and evaluation of interventions. *International Journal of Behavioral Development*, 34(5): 473-480.

Disclosures

KRN, CKR, and KM are employees of INC Research, the Contract Research Organization responsible for execution of all trials included in this analysis.

Poster presentation at the 2015 Autumn International Society of Clinical Trials Methodology (ISCTM) Conference, Amsterdam, 27-29 August 2015

inc Research®