



Published in final edited form as:

Drug Alcohol Depend. 2013 February 1; 128(1-2): 77–82. doi:10.1016/j.drugalcdep.2012.08.007.

Assessment of Riboflavin as a Tracer Substance: Comparison of a Qualitative to a Quantitative Method of Riboflavin Measurement

Abigail J. Herron, D.O.,

St. Luke's Roosevelt Hospital Center, 324 W. 108th Street, #505, New York, NY 10025, Tel: 212-280-0123, Fax: 212-864-6380, aherron@chpnet.org, aherron576@gmail.com

John J. Mariani, M.D.,

New York State Psychiatric Institute, 1051 Riverside Drive, Unit 66, New York, NY 10032

Martina Pavlicova, Ph.D.,

New York State Psychiatric Institute, 1051 Riverside Drive, Unit 51, New York, NY 10032

Christina M. Parinello, M.P.H.,

Albert Einstein College of Medicine, Department of Epidemiology and Population Health, 1300 Morris Park Ave., Belfer 1306B, Bronx NY 10461

Krysten W. Bold, B.S.,

Rutgers, The State University of New Jersey, 152 Frelinghuysen Road, Piscataway NJ 08854

Frances R. Levin, M.D.,

New York State Psychiatric Institute, 1051 Riverside Drive, Unit 66, New York, NY 10032

Edward V. Nunes, M.D.,

New York State Psychiatric Institute, 1051 Riverside Drive, Unit 120, New York, NY 10032

Maria A. Sullivan, M.D., Ph.D.,

New York State Psychiatric Institute, 1051 Riverside Drive, Unit 66, New York, NY 10032

Wilfred N. Raby, M.D., Ph.D., and

New York State Psychiatric Institute, 1051 Riverside Drive, Unit 120, New York, NY 10032

Adam Bisaga, M.D.

New York State Psychiatric Institute, 1051 Riverside Drive, Unit 66, New York, NY 10032

Abstract

Background—Noncompliance with medications may have major impacts on outcomes measured in research, potentially distorting the validity of controlled clinical trials. Riboflavin is frequently used in trials as a marker of adherence. It can be combined with study medication and is excreted in urine where it fluoresces under UV light. This study compares qualitative visual inspection of fluorescence to quantitative fluorometric analysis of riboflavin concentration in its ability to detect the presence of riboflavin in urine.

Methods—Twenty-four volunteers received 0 mg, 25 mg, and 50 mg doses of riboflavin under single-blind conditions, with 20 also receiving a 100 mg dose. Five serial urine samples were collected over the following 36 hours. Quantitative measurement of riboflavin by fluorometric analysis and qualitative assessment of each sample using visual inspection were performed.

Results—The overall false positive rate for qualitative assessment was 53%. For quantitative assessment, a riboflavin concentration of 900 ng/mL was established to classify positive samples. More than 80% of samples were positive 2 to 24 hours following ingestion of 25 mg and 50 mg,

and less than 80% were positive at 36 hours. At least 95% of observations for the 100 mg dose were above 900 ng/mL at all timepoints.

Conclusions—Quantitative fluorometric assessment is superior to qualitative visual inspection alone in determining medication adherence. The combination of 25–50 mg of daily riboflavin and a cut-off level of 900 ng/mL allows for the acceptable sensitivity of missing detection of non-compliant participants while preserving a high level of power to detect all cases of medication compliance.

Keywords

adherence; quantitative; qualitative; riboflavin; tracer

1. INTRODUCTION

Non-adherence with medications threatens the validity of controlled clinical trials. Assessment of medication adherence is essential to determining whether the study has provided an adequate evaluation of the study medication (Babiker et al., 1986; Kranzler et al., 1997). Similarly, poor adherence can compromise treatment and lead to poor clinical outcomes.

A number of variables have been identified as contributing to nonadherence, including medication factors, such as dosing frequency, cost, or side effects; illness factors; and patient factors, such as cognitive impairment and comorbid psychiatric illness (Osterburg and Blaschke, 2005; Swift, et al., 2011). Individuals with substance use disorders experience many of these, contributing to a particular problem with adherence in this population. Substance use can directly impair judgment, thereby negatively impacting treatment adherence (Magura, et al., 2002). Additionally, there can be ambivalence surrounding the use of medications in this population, which can affect adherence (Sowers and Golden, 1999). In a study of 577 individuals receiving disulfiram for alcohol dependence, Fuller et al. (1986) found only 20% were adherent to medication. Teter et al. (2011) studied individuals with bipolar disorder and found that rates of medication adherence were lower in individuals with current substance use disorders as compared to those with past history of or no substance disorder.

Indirect methods for assessing adherence, such as patient self-report, doctors' estimates of adherence, pill counts and specialized packaging have been shown to be less reliable than direct methods such as measurement of drug levels, (Roth and Caron, 1978; Epstein and Cluss, 1982; Kranzler et al., 1997), and these methods may result in overestimates of compliance (Spilker, 1991). Pill counts may provide a better assessment of adherence than self-reports, but depend on the premise that all missing pills were taken by the patient, when they may have been removed from the packaging and subsequently discarded or shared with another individual. Tests of the validity of pill counts have yielded equivocal results (Roth et al., 1970; Epstein and Cluss, 1982; Pullar et al., 1989; Rudd et al., 1989; Young et al., 1984). Micro-electronic monitoring, such as containers that record the time and frequency of opening, provide continuous assessments of compliance, but can be costly and inaccurate. Patients may remove more than one pill at each opening, or not ingest all pills removed from the container (Cheung et al., 1988; Cramer et al., 1989; Rudd et al., 1981). Measurement of drugs or their metabolites in body fluids allows for a more direct assessment of adherence. This method, however, requires special laboratory facilities and can be expensive. Additionally, the interpretation of results can be confounded by differences in drug metabolism and cumulative effects of long-acting compounds. More importantly, the

measure of compliance is applied to the active medication group but not to the placebo group, which limits its usefulness.

Tracer assay, in which a compound is encapsulated with study medication and is later detectable in urine, provides an alternative to directly measuring drug levels. A number of substances have been employed as tracers, including sodium bromide, methylene blue, phenol red, fluorescein, bromide, and phenazopyridine (Roth et al., 1970; Epstein and Masek, 1978; Spriet and Simon, 1985; Kraus et al., 1987), but riboflavin has been the most extensively studied (Hobby and Deuschle, 1959; Deuschle et al., 1960; Veterans Administration Cooperative Study Group, 1967, 1970, 1977a,b; Cluss and Epstein, 1984; Cluss et al., 1984; Dubbert et al., 1985; Del Boca, et al., 1996). It is a nontoxic, non-allergenic, water-soluble vitamin (B2), which is inexpensive and readily absorbed through oral administration. Riboflavin is excreted in urine at relatively low concentrations, where it can be detected by fluorescence under UV light using quantitative fluorometric assessment or qualitative visual inspection (Hobby and Deuschle, 1959; Spilker, 1991; Lambert et al., 1985). After administration, riboflavin is rapidly absorbed (t_{\max} 1.4–2 hours) and is eliminated in urine, with more than 91% of the total excretion of riboflavin taking place during the first 24 hours (Zempleni et al., 1996), making it a good candidate for the measurement of compliance using a once per day dosing.

The presence of riboflavin in many foods and vitamin products is a potential disadvantage to its use as a tracer, as dietary riboflavin results in low, background levels of UV fluorescence. However, urine riboflavin levels are 10–20 times higher following the ingestion of supplementary riboflavin at the 20–60 mg dose (Malcolm et al., 1992; Zempleni et al., 1996). With large doses of riboflavin or multiple daily doses, a “spillover” effect has been observed, in that riboflavin can be detected in the urine for longer than 24 hours (Babiker et al., 1989; Cluss and Epstein, 1984). This is a significant limitation of the utility of riboflavin in cases of multiple-daily dosing. Riboflavin has been reported to be light-sensitive (Chen et al., 2005), and a number of studies have kept specimens refrigerated and protected from sunlight. Babiker et al. (1989), however, demonstrated no difference in fluorescence among specimens that were refrigerated, those kept at room temperature but protected from light, and specimens kept at room temperature and exposed to light.

Studies regarding the reliability of riboflavin detection have yielded equivocal results. Anton (1996) assessed the accuracy of visual inspection using results from fluorometric analysis as a standard and found an overall error rate of 28%, comprised of a 20% false positive rate and an 8% false negative rate, suggesting that qualitative assessment is less precise than quantitative fluorometry. The VA Cooperative Group (Goldman et al., 1982) found detection of riboflavin fluorescence to be unreliable, dropping this measure from their analysis of medication compliance because of lack of uniform interpretations of fluorescence using visual inspection. Young et al. (1984) also report that riboflavin testing was discontinued in previous VA Cooperative Group studies because of difficulty in detection by the clinician.

Other studies, however, have found that riboflavin fluorescence could be accurately detected even by individuals with minimal training (Cluss and Epstein, 1984; Cluss et al., 1984). Dubbert et al. (1985) used visual inspection to differentiate fluorescence of normal dietary riboflavin from that produced by riboflavin supplementation. The authors found that the accuracy of discrimination increased with riboflavin dose, and reached 100% with measurement 2 to 8 hours after a 50 mg dose, although there was an average 21.2% false positive rate. Babiker (1989) also demonstrated that a single 50 mg dose could be detected via visual inspection within 4 to 6 hours with 95% accuracy in a small study of five participants.

Del Boca et al. (1996) performed six trials in which they assessed the reliability and validity of visual inspection, and determined that raters were most reliable in correctly identifying negative samples and those which contained high concentrations of riboflavin. They used additional measures for validity—employing diaries of pills taken in two of the trials, a capsule count in one trial, and fluorometric analysis in another, and demonstrated accuracy of up to 94%. This study was limited significantly by small sample size, however. Five of the six trials had only two participants, including the one using fluorometric validation of results.

Many research groups use a qualitative assessment of fluorescence, but we were not able to locate studies that have verified the results of qualitative assessment using objective measurements to provide a decisive assessment of the validity of visual inspection. To address this, we compared results of qualitative visual inspection to quantitative fluorometric measurement of riboflavin concentration. We sought to determine whether the qualitative method has acceptable performance to be used as a method of monitoring medication compliance and to determine the dose of riboflavin that is optimal for monitoring medication compliance using the quantitative method of measurement.

2. METHODS

2.1 Participants

The sample consisted of twenty-four healthy individuals recruited through word-of-mouth and posted flyers. Eligible participants were between the ages of 18 and 50, willing to participate and comply with study procedures, and classified as medically healthy volunteers after a general medical history interview. The study was approved by the New York State Psychiatric Institute Institutional Review Board. All study participants provided written informed consent and were compensated for taking part in the study. Individuals who reported current renal or urinary tract dysfunction, pregnancy or lactation, or a history of allergic reaction to riboflavin were excluded from participation.

2.2 Procedures

Riboflavin capsules (0 mg, 25 mg, 50 mg and 100 mg) were administered in single doses at least one week apart. Twenty-four participants received the 0 mg, 25 mg, and 50 mg doses, and twenty of those participants also received the 100 mg dose. In each dosing phase, the riboflavin capsule was administered under single-blind counter-balanced conditions (0 h) and five serial urine samples were collected over the following 36 hours (2, 6, 8, 24 and 36 hours). At the time of riboflavin administration, participants received 5 tubes for specimen collection, as well as instructions with specific times when urine samples should be collected. Participants were reminded by phone and/or email when specimens should be collected throughout the day.

No dietary conditions were imposed, but to reduce the potential confound of exogenous riboflavin ingestion, subjects were instructed to avoid consumption of all vitamin supplements for at least one week before participating in the study. Urine specimens were refrigerated and protected from direct light. All participants were provided with darkly colored bags for specimen storage and transportation. Specimens were kept under identical conditions prior to qualitative and quantitative assessments.

2.3 Measures

2.3.1 Qualitative Assessment—The qualitative measurement of riboflavin was performed by three independent raters who were blind to dose condition. Raters visually inspected the collected specimens for presence or absence of fluorescence after nonspecific

ultraviolet (UV) light activation in a dark room. Each rater was trained and experienced with this methodology for at least one year in clinical trials of drug addiction pharmacotherapies. All samples were examined twice by all three raters. At any one rating session, raters made approximately 50 assessments.

2.3.2 Quantitative Assessment—The quantitative measurement of riboflavin was performed by assaying riboflavin directly against a linear standard curve using a TD-700 laboratory fluorometer (Turner Designs, Sunnyvale CA). The fluorometer was calibrated using a riboflavin stock solution diluted with a sodium acetate buffer to yield five known concentrations, ranging from 80 ng/mL to 800 ng/mL (Henderleiter and Hyslop, 1996). Once calibrated, other samples of known concentration were analyzed to verify the accuracy of the digital readout. All urine specimens were excited at 450 nm and their emission fluorescence measured and recorded at 525 nm. Highly concentrated urine samples were diluted using distilled water to yield concentrations within the linear standard curve range. The need for dilution was dose-dependent. At the 0 mg dose, 2.5% (3 of 120) of the samples required dilution; at 25 mg, 88.3% (106 of 120 samples); at 50 mg, 91.6% (110 of 120); and at 100 mg, 100% (100 of 100). All samples were fully analyzed within 72 hours of specimen collection.

2.4 Statistical Analysis Plan

As the qualitative test is mostly used to identify non-compliant participants, we first wanted to assess the likelihood that the test would overestimate the rates of compliance. We calculated the false positive rate, defined as the percentage of observations rated as positive for presence of riboflavin in the 0 mg dose condition. We also calculated Cohen's Kappa separately for each rater to assess agreement between each individual's rating of the 0 mg dose. Two ratings of presence of riboflavin or two ratings of absence of riboflavin for the same sample were considered to be concordant ratings, while one rating of each was considered to be a discordant rating.

In the next stage of analysis, we measured the concentration of riboflavin for all collected samples. In order to determine the rate of false positives for the quantitative method, subjects were randomly separated into two equal groups. The 95th percentiles of riboflavin concentration in urine after ingestion of 0 mg of riboflavin were determined using one group (and defined as a cut-off) and tested using the second group. Subjects who were given the 0 mg dose but tested above this determined cut-off were considered to have false positive results. Next, we applied a cut-off value to the results of each riboflavin dose at each time point to determine the true positive rate (%) for that particular cut-off.

Finally, we wanted to determine the accuracy of the qualitative method to detect fluorescence positive samples. We calculated Pearson correlation coefficients for each time point and dose to determine how well negative or positive visual assessments correlated with negative or positive quantitative assessments. Both qualitative and quantitative results were categorized as binary variables.

For the qualitative assessment, an absence of fluorescence was considered to be a negative result, and presence of fluorescence was a positive result. For the quantitative assessment, riboflavin concentration equal to or lower than the predetermined cut-off value was considered a negative result, while a reading greater than the cut-off point was considered a positive result.

Analyses were conducted using SAS (version 9.2; Cary, NC).

3. RESULTS

3.1 Study participants

Twenty-four volunteers took part in the study, with a mean age of 27 years (SD=5.5). The sample was 67% female (n=16), 67% Caucasian (n=16), 8% African-American (n=2), 4% Hispanic (n=1), and 21% Asian (n=5).

3.2 Qualitative assessment

The overall false positive rate for the visual assessment of the 0 mg riboflavin dose was 53% (Table 1). Cohen's kappa to measure agreement between each rater's two assessments for each participant ranged from 0.35 to 0.69.

3.3 Quantitative assessment

The concentration of riboflavin in 0 mg samples varied between 80 and 3,879 ng/mL. The riboflavin concentration cut-off value to classify samples as negative, determined from the distribution of the riboflavin concentration in all samples from the 0 mg condition in the first group, was set at 900 ng/mL (approximately the 95th percentile). Using this cutoff, we classified samples from two (3.3%) subjects in the first group as false positives (above 900 ng/mL). When this cut-off value was applied to results from the second group, samples from four participants (6.7%) were false positives. Therefore, we accepted 900 ng/mL as a reasonable cut-off, resulting in about a 5% false positive rate. We have used the determined values to plot changes in the concentration of riboflavin over time in all four groups (Figure 1).

Subsequently, we applied the cut-off of 900 ng/mL to all samples for which riboflavin concentration was determined, resulting in more than 80% of samples above 900 ng/mL at 25 mg and 50 mg from 2 to 24 hours, and less than 80% above the cut-off at 36 hours. At least 95% of observations for the 100 mg dose were above 900 ng/mL at all time points (Figure 2).

3.4 Comparison of qualitative and quantitative assessments

Comparison of the accuracy of visual assessment (positive or negative fluorescence) with the quantitative assessments (above or below a 900 ng/ml threshold) showed that there was no significant correlation at the 0 mg dose. However, the mean for the group classified as positive on the majority of qualitative assessments was significantly higher than the mean for the group classified as mostly negative (576 vs. 392 ng/ml, $p=0.04$).

There was a high variability for the 25 mg and 50 mg doses. Qualitative and quantitative assessments were highly correlated for the 100 mg dose ($P<0.01$) at hours 6 and 8. Correlations were indeterminable at hours 2 and 24 because either all quantitative readings were above 900 ng/mL or all rater responses were positive. Visual and quantitative assessments were not significantly correlated at hour 36.

4. DISCUSSION

Our results demonstrate that the visual inspection method of detecting riboflavin fluorescence is highly unreliable at correctly classifying urine samples as riboflavin-negative in participants who do not take supplemental, non-dietary riboflavin. Additionally, the accuracy of classifying urine samples using visual inspection was shown to be highly variable among raters.

Testing a range of urine samples known to have no supplemental riboflavin, we determined a level of 900 ng/mL to be a reasonable cut-off value for quantitative analysis, above which samples can be considered positive for the ingestion of supplemental riboflavin. With this cut-off, the quantitative method allows the detection of riboflavin for up to 24 hours following the ingestion of 25 or 50 mg of riboflavin, with a power of more than 80%. At the dose of 100 mg, concentrations of riboflavin remain elevated above the cut-off level beyond the 24-hour time-point.

With larger doses, the level of riboflavin fails to return to baseline within the subsequent 24-hour, producing the spillover effect previously reported (Babiker et al., 1989; Cluss and Epstein, 1984). This finding supports the use of a 25 mg or 50 mg dose administered once per day in the context of a clinical trial in order to maximize the detection of true positive samples, representing patients compliant with study medications while minimizing the possibility of false-positive extending beyond a 24-hour time frame, representing non-compliant patients. With a dose of 100 mg and a cut-off of 900 ng/mL the method is well suited to maximize the correct detection of all participants who are compliant with medication when detecting noncompliant individuals is less of a priority. On the other hand, using a dose of 25 mg will maximize the detection of non-compliant individuals while sacrificing the power to detect compliant individuals. In situations where the detection of most instances of non-compliance is a priority, the low riboflavin dose and a higher cut-off value for the negative samples (e.g., to 2,000 ng/ml) should be utilized. We suggest that higher daily doses of riboflavin are appropriate to use in testing compliance with longer-acting medications, while lower doses will allow the most accurate detection of participants who miss doses, and are appropriate to use in testing compliance with shorter-acting medications (dosed more than once daily).

The high error rate of the qualitative method demonstrated here is consistent with results of studies by Anton (1996) and Dubbert et al. (1985), although we found the rate of false positives to be even higher. Visual inspection was shown to be highly unreliable with variable accuracy among raters, consistent with findings by the VA Cooperative Group, who discontinued riboflavin testing as a result. This is due in part to the presence of riboflavin in several food groups (e.g., coffee, red meat, cereals), multivitamin preparations, and as an additive to food supplements and colorants (e.g., energy drinks, orange-flavor drinks). Accuracy improved with increased dose of riboflavin, an observation which has also been reported in previous trials (Dubbert, et al., 1985; Del Boca, 1996).

Assessment of medication adherence using visual inspection alone likely results in a significant overestimation of medication compliance in clinical trials, given the high rate of false-positive samples demonstrated here. Even with the trained and experienced staff, the detection of negative samples was little better than at the chance level, and therefore we cannot recommend this method for use in clinical trials. The quantitative assessment method can provide a more accurate assessment of medication compliance in clinical trials. For that, a supplemental riboflavin needs to be encapsulated with a study drug or placebo, with a total dose of 25–50 mg per day, which permits some flexibility when a dose of study drug needs to be adjusted. The quantitative method is more reliable than the qualitative method, as it eliminates the difference between raters' subjective assessments, which can be significant. However this method requires the added time for sample preparation and testing as well as specialized equipment.

The relatively long time to return to negative levels following ingestion of riboflavin makes the fluorometric analysis useful primarily for once-daily dosing regimens, which is a significant limitation of its use. Additionally, detection of riboflavin in urine measures

discrete adherence events over a short time period (24 hours) and therefore is best suited to studies that include frequent clinic visits.

In the setting of a clinical trial to assess the effectiveness of a medication, the frequent assessment of medication compliance is most helpful to identify participants that consistently do not take study medication, and allows for the possibility of reanalyzing the primary outcome with non-compliant individuals removed. Assessment of non-compliance can also be used to identify, and intervene with, participants who have variable medication compliance. However the riboflavin method, with its limited ability to detect occasional non-compliance, is less suited for that purpose. In that case a marker with a long “half-life” that will reflect the compliance averaged over the several days will be more appropriate.

This study has several limitations. While participants were reminded via text message and/or email when specimens were due to be obtained, the collection of urine specimens was not directly observed. We cannot be certain that all urine specimens were collected at the specified timepoints. During the visual assessment task, raters were instructed only to rate samples as positive or negative. Useful information may have been obtained if responses such as “strongly positive, weakly positive, or unsure” had been included. Additionally, raters may have been inclined to over-report samples as positive because of a belief that the research participants had been adherent with the study design. More detailed instructions may improve the usefulness of the method. While the 25 mg and 50 mg doses were shown to produce positive readings up to 24 hours, the timing of the readings left a large window of time between 8 and 24 hours after ingestion without any assessments. It would have been useful to determine the pattern of fluorescence during that time period. Finally, we have assumed that a single-dose paradigm used in this study can model the clinical situation of patients taking medication daily, on a chronic basis. While the majority of riboflavin exogenous riboflavin is excreted within the 24 hours after the dose, it is possible that there may be some additional accumulation with chronic dosing which may contribute to even greater “spillover effect” than observed in our study.

In summary, our results indicate that quantitative fluorometric assessment is superior to qualitative visual inspection alone in determining riboflavin-labeled medication adherence. The combination of 20–50 mg of daily riboflavin dose and a cut-off level of 900 ng/mL allows for the acceptable sensitivity of missing detection of non-compliant participants while preserving a high level of power to detect all compliance with the medication.

REFERENCES

- Anton RF. New Methodologies for Pharmacological Treatment Trials for Alcohol Dependence. *Alcohol Clin Exp Res.* 1996; 20:3A–9A. [PubMed: 8651457]
- Babiker IE. Noncompliance in schizophrenia. *Psychiat Dev.* 1986; 4:329–337.
- Babiker IE, Cooke PR, Gillett MG. How Useful is Riboflavin as a Tracer of Medication Compliance? *J Behav Med.* 1989; 12:25–38. [PubMed: 2746641]
- Chen M, Andrenyak DM, Moody DE, Foltz RL. Determination of riboflavin by high-performance liquid chromatography with riboflavin-depleted urine as calibration and control matrix. *J Chrom B.* 2005; 820:147–150.
- Cheung R, Dickins J, Nicholson PW, Thomas ASC, Smith HH, Larson HE, Deshmukh AA, Dobbs RJ, Dobbs SM. Compliance with anti-tuberculosis therapy: A field trial of a pill-box with a concealed electronic recording device. *Eur J Clin Pharmacol.* 1988; 35:401–407. [PubMed: 3197749]
- Cluss PA, Epstein LH. A riboflavin tracer method for assessment of medication compliance in children. *Behav Res Meth Ins C.* 1984; 16:444–446.
- Cluss PA, Epstein LH, Galvis SA, Fireman P, Friday G. Effect of compliance for chronic asthmatic children. *J Consult Clin Psych.* 1984; 52:909–910.

- Cramer JA, Mattson RH, Prevey ML, Scheyer RD, Ouellette VL. How often is medication taken as prescribed? A novel assessment technique. *JAMA*. 1989; 261:3273–3277. [PubMed: 2716163]
- Del Boca FK, Kranzler HR, Brown J, Korner PF. Assessment of medication compliance in alcoholics through UV light detection of a riboflavin tracer. *Alcohol Clin Exp Res*. 1996; 20:1412–1417. [PubMed: 8947318]
- Deuschle KW, Jordahl C, Hobby GL. Clinical usefulness of riboflavin-tagged isoniazid for self-medication in tuberculosis patients. *Am Rev Respir Dis*. 1960; 82:1–10. [PubMed: 13816252]
- Dubbert PM, King A, Rapp SR, Brief D, Martin JE, Lake M. Riboflavin as a tracer of medication compliance. *J Behav Med*. 1985; 8:287–299. [PubMed: 4087291]
- Epstein LH, Cluss PA. A behavioral medicine perspective on adherence to long-term medical regimens. *J Consult Clin Psych*. 1982; 50:950–971.
- Epstein LH, Masek BJ. Behavioral control of medicine compliance. *J Appl Behav Anal*. 1978; 11:1–9. [PubMed: 649520]
- Fuller RK, Branchey L, Brightwell DR, Derman RM, Emrick CD, Iber FL, James KE, Lacoursiere RB, Lee KK, Lowenstam I, Maany I, Neiderhiser D, Nocks JJ, Shaw S. Disulfiram Treatment of Alcoholism. A Veterans Administration Cooperative Study. *JAMA*. 1986; 256:1449–1455. [PubMed: 3528541]
- Goldman AI, Holcomb R, Perry HM Jr, Schnaper W, Fitz AE, Frohlich ED. Can dropout and other noncompliance be minimized in clinical trials? *Control Clin Trials*. 1982; 3:75–89. [PubMed: 6749426]
- Henderleiter JA, Hyslop RM. The Analysis of Riboflavin in Urine Using Fluorescence. *J Chem Educ*. 1996; 73:563–565.
- Hobby GL, Deuschle KW. The use of riboflavin as an indicator of isoniazid ingestion in self-medicated patients. *Am Rev Respir Dis*. 1959; 80:415–423.
- Kranzler, HR.; Mason, BJ.; Pettinati, HM.; Anton, RF. Methodological issues in pharmacotherapy trials with alcoholics. In: Hertzman, M.; Feltner, D., editors. *The Handbook of Psychopharmacology Trials*. New York: New York University Press; 1997. p. 213-245.
- Kraus RP, Grof P, Arana GW, Workman RJ, Harvey KJ, Hux M. Methylene blue: A reliable and practical marker for validating compliance on the DST. *J Clin Psychiat*. 1987; 48:224–229.
- Lambert WE, Cammaert PM, De Leenheer AP. Liquid-chromatographic measurement of riboflavin in serum and urine with isoriboflavin as internal standard. *Clin Chem*. 1985; 31:1371–1373. [PubMed: 4017244]
- Magura S, Laudet AB, Mahmood D, Rosenblum A, Knight E. Adherence to Medication Regimens and Participation in Dual-Focus Self-Help Groups. *Psychiatr Serv*. 2002; 53:310–316. [PubMed: 11875225]
- Malcolm R, Anton RF, Randall CL, Johnston A, Brady K, Thevos A. A placebo-controlled trial of buspirone in anxious inpatient alcoholics. *Alcohol Clin Exp Res*. 1992; 16:1007–1013. [PubMed: 1335217]
- Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med*. 2005; 353:487–497. [PubMed: 16079372]
- Pullar T, Kumar S, Tindall H, Feely M. Time to stop counting the tablets? *Clin Pharmacol Ther*. 1989; 46:163–168. [PubMed: 2758726]
- Roth HP, Caron HS. Accuracy of doctors' estimates and patients' statements on adherence to a drug regimen. *Clin. Pharmacol. Ther*. 1978; 23:361–370. [PubMed: 627144]
- Roth HP, Caron HS, Hsi BP. Measuring intake of a prescribed medication: A bottle count and a tracer technique compared. *Clin. Pharmacol. Ther*. 1970; 11:228–237. [PubMed: 5416853]
- Rudd P, Byyny RL, Zachary V, LoVerde ME, Titus C, Mitchell WD, Marshall G. The natural history of medication compliance in a drug trial: Limitations of pill counts. *Clin Pharmacol Ther*. 1989; 46:169–176. [PubMed: 2667837]
- Rudd P, Marshall G, Taylor CB, Agras WS. Medication monitor/dispenser for pharmaceutical and compliance research. *Clin Pharmacol Ther*. 1981; 29:278.
- Sowers W, Golden S. Psychotropic medication management in persons with co-occurring psychiatric and substance use disorders. *J Psychoactive Drugs*. 1999; 31:59–70. [PubMed: 10332640]

- Spielher V, Fay J, Fogerson R, Schoendorfer D, Niedbala RS. Enzyme immunoassay validation for qualitative detection of cocaine in sweat. *Clin Chem*. 1996; 42:34–38. [PubMed: 8565229]
- Spilker, B. *Guide to Clinical Trials*. New York: Raven Press; 1991. p. 102-114.
- Spriet, A.; Simon, P. *Methodology of Clinical Drug Trials*. Edelstein, R.; Weintraub, M., translators. Basel, Switzerland: Karger Publishers; 1985. p. 127-135.
- Swift R, Oslin DW, Alexander M, Forman R. Adherence monitoring in naltrexone pharmacotherapy trials: a systematic review. *J Stud Alcohol Drugs*. 2011; 72:1012–1018. [PubMed: 22051215]
- Teter CJ, Falone AE, Bakaian AM, Tu C, Ongür D, Weiss RD. Medication adherence and attitudes in patients with bipolar disorder and current versus past substance use disorder. *Psychiatry Res*. 2011; 190:253–258. [PubMed: 21696830]
- Veterans Administration Cooperative Study Group. Effects of treatment on morbidity in hypertension: Results in patients with diastolic blood pressures averaging 115 through 129 mm Hg. *JAMA*. 1967; 202:1028–1034. [PubMed: 4862069]
- Veterans Administration Cooperative Study Group. Effects of treatment on morbidity in hypertension: Results in patients with diastolic blood pressure averaging 90 through 114 mm Hg. *JAMA*. 1970; 213:1143–1152. [PubMed: 4914579]
- Veterans Administration Cooperative Study Group. Propranolol in the treatment of essential hypertension. *JAMA*. 1977a; 237:2303–2310. [PubMed: 323525]
- Veterans Administration Cooperative Study Group. Multiclinic controlled trial of bethanidine and guanethidine in severe hypertension. *Circulation*. 1977b; 55:519–525. [PubMed: 319922]
- Weiss RD. Adherence to pharmacotherapy in patients with alcohol and opioid dependence. *Addiction*. 2004; 99:1883–1892.
- Young LM, Haakenson CM, Lee KK, van Eeckhout JP. Riboflavin Use as a Drug Marker in Veterans Administration Cooperative Studies. *Control Clin Trials*. 1984; 5:497–504. [PubMed: 6394211]
- Zempleni J, Galloway JR, McCormick DB. Pharmacokinetics of orally and intravenously administered riboflavin in healthy humans. *Am J Clin Nutr*. 1996; 63:54–66. [PubMed: 8604671]

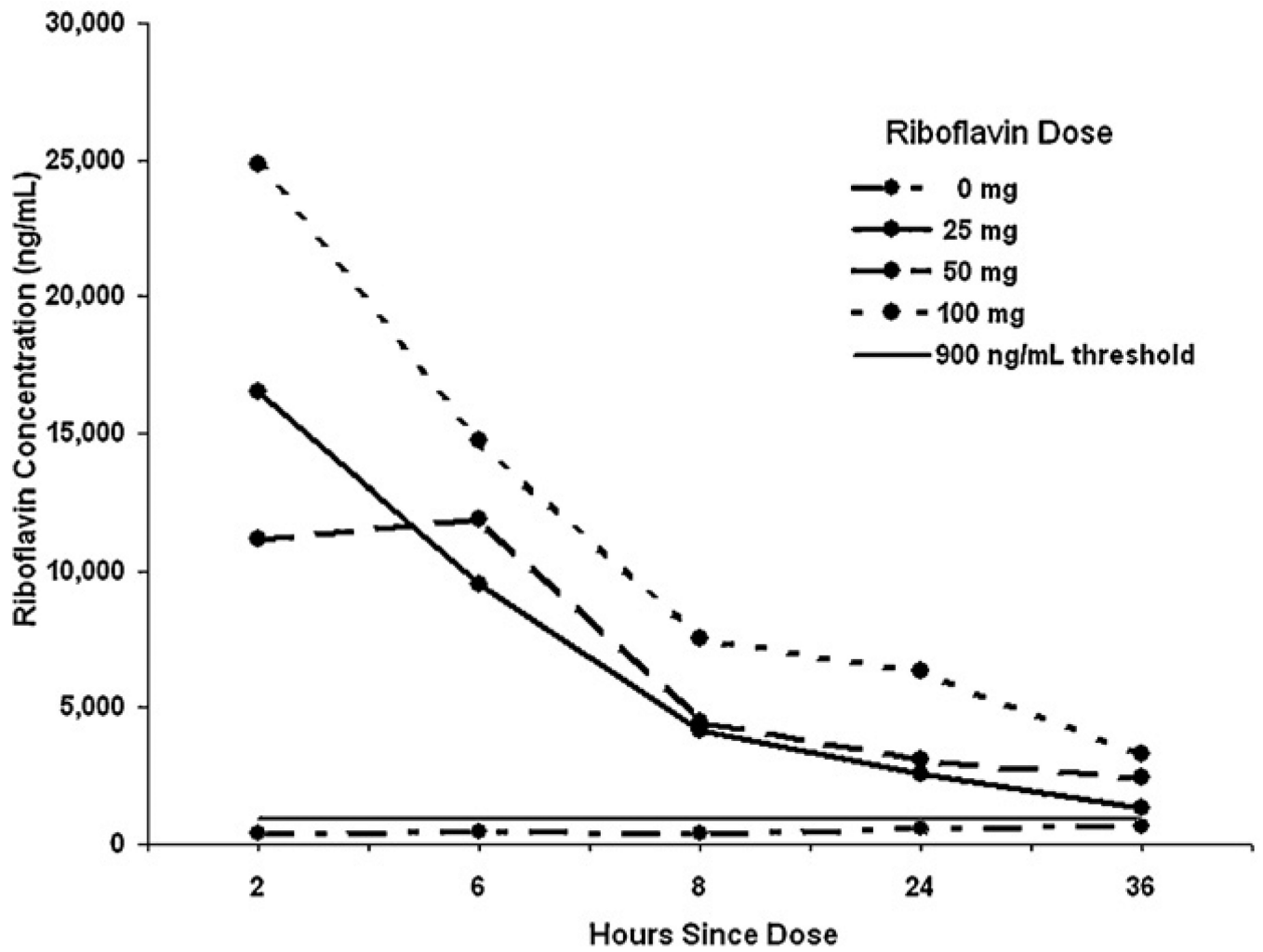


Figure 1. Concentration of riboflavin (mean \pm SE) in the urine at various timepoints following ingestion of capsules containing riboflavin 0–100 mg (N=24).

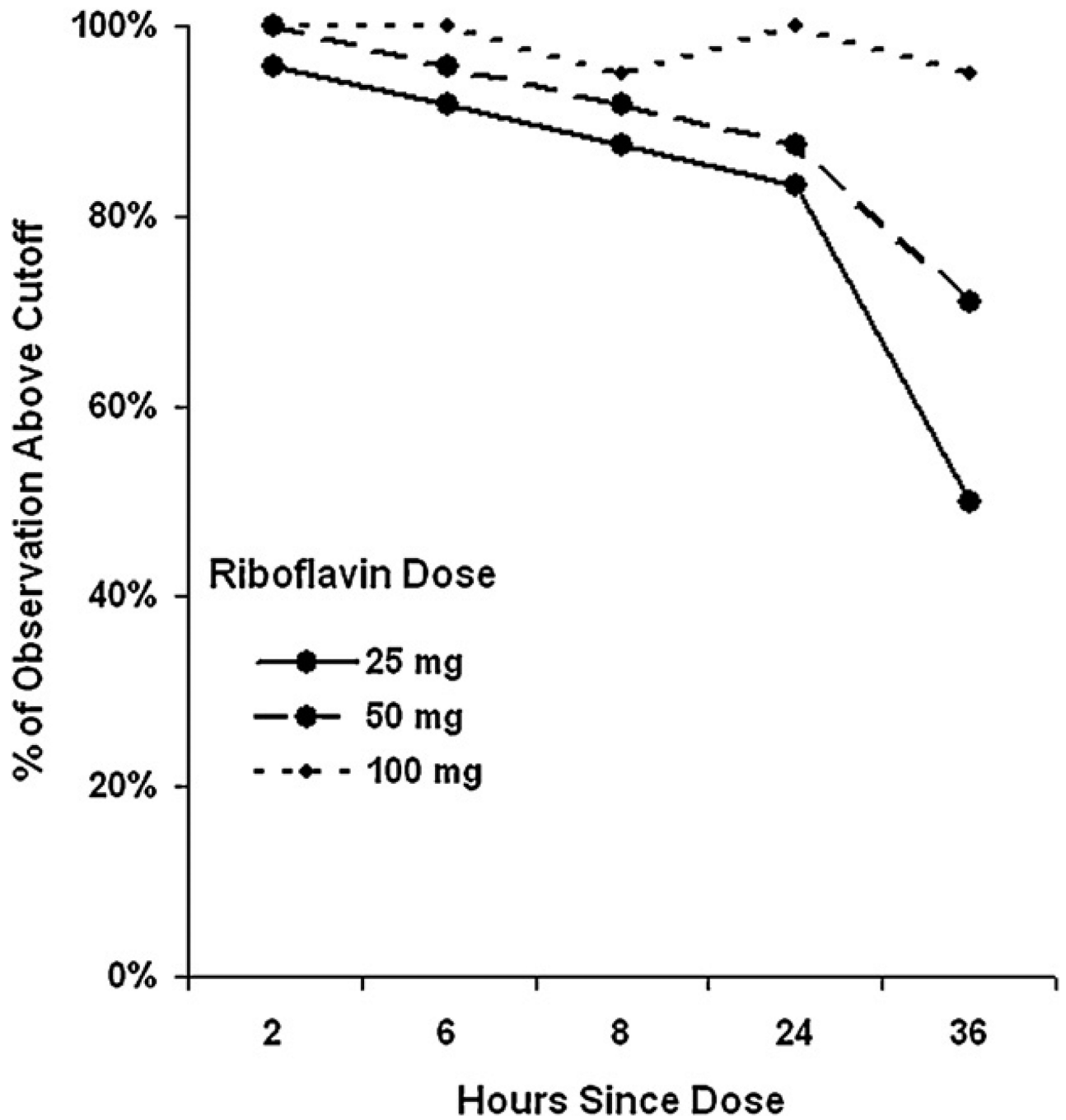


Figure 2. Ratio of urine samples with detected riboflavin concentration above the cut-off value at various timepoints following ingestion of capsules containing riboflavin 0–100 mg (N=24).

Table 1

False Positives and Specificity for the Qualitative Visual Inspection at the 0 mg dose

	False Positives	Specificity*
Rater A	55.67%	43.33%
Rater B	55.42%	44.58%
Rater C	48.33%	51.67%
Overall	53.5%	46.5%

* Specificity = (True Negatives / (False Positives + True Negatives))