

---

# Opioid Pharmacogenomics Using a Twin Study Paradigm: Methods and Procedures for Determining Familial Aggregation and Heritability

Martin S. Angst,<sup>1</sup> Nicholas G. Phillips,<sup>1</sup> David R. Drover,<sup>1</sup> Martha Tingle,<sup>1</sup> Jeffrey L. Galinkin,<sup>2</sup> Uwe Christians,<sup>2</sup> Gary E. Swan,<sup>3</sup> Laura C. Lazzeroni,<sup>4</sup> and J. David Clark<sup>1,5</sup>

<sup>1</sup> Department of Anesthesia, Stanford University School of Medicine, United States of America

<sup>2</sup> Department of Anesthesiology, Children's Hospital, University of Colorado at Denver Health Science Center, Aurora, United States of America

<sup>3</sup> Center for Health Sciences, SRI International, Menlo Park, CA, United States of America

<sup>4</sup> Department of Psychiatry and Behavioral Sciences and of Pediatrics, Stanford University School of Medicine, United States of America

<sup>5</sup> Department of Anesthesia, Veterans Affairs Palo Alto Health Care System, Palo Alto, CA, United States of America

Opioids are the cornerstone medication for the treatment of moderate to severe pain. However, analgesic opioid requirements and the propensity to suffer from aversive opioid effects, including fatal respiratory depression and addiction, vary widely among patients. The factors underlying the substantial response variance remain largely unknown and need clarification for using opioids more effectively in appropriately selected patients. This ongoing study takes advantage of the twin paradigm to estimate the genetic and environmental contributions to inter-individual differences in opioid responses. Evidence of significant heritability will justify more detailed and extensive genomic studies. The enrollment target is 80 monozygotic and 45 dizygotic twin pairs who undergo a target-controlled infusion of the opioid alfentanil and saline placebo in sequential but randomized order. In a laboratory-type setting, well-defined pharmacodynamic endpoints are measured to quantify pain sensitivity, analgesic opioid effects, and aversive opioid effects including respiratory depression, sedation and reinforcing affective responses. First results obtained in 159 participants provide evidence for the feasibility and utility of this interventional study paradigm to estimate familial aggregation and heritability components of relevant drug effects. Areas highlighted in this report include recruitment strategies, required infrastructure and personnel, selection of relevant outcome measures, drug infusion algorithm minimizing pharmacokinetic variability, and considerations for optimizing data quality and quantity without hampering feasibility. Applying the twin paradigm to complex and potentially harmful studies comprehensively characterizing pharmacological response profiles is without much precedent. Methods and first results including heritability estimates for heat and cold pain sensitivity should be of interest to investigators considering similar studies.

**Keywords:** pharmacogenomics, pharmacodynamics, opioid, alfentanil, twin, heritability, familial aggregation, heat pain, cold pressor pain, pain sensitivity, quantitative sensory test, analgesia, respiratory depression, sedation, nausea, pruritus, liking, positive affective response

---

The experience of pain is a normal part of life, and the absence of the ability to experience pain is associated with severe problems, most notably the inability to perceive self-injury (Clark, 2008). However, unrelieved pain is a vexing clinical problem as well. For example, despite aggressive management, moderate to severe pain is experienced by 20 to 40% of patients after surgery (Dolin et al., 2002; Michel & Sanders, 2003). Even though several treatment options are available including topical agents, nonsteroidal anti-inflammatory drugs, and disease-modifying drugs for diseases such as rheumatoid arthritis, a minority of these patients have levels of pain they consider 'acceptable' (Heiberg et al., 2005; Heiberg & Kvien, 2002). Chronic pain is a highly prevalent condition experienced by more than 30% of children and adults (Bouhassira et al., 2008; Clark, 2002). The overall costs of chronic pain to the US economy exceed \$60 billion annually and are caused by medical evaluations, medications, procedures and lost productivity (Stewart et al., 2003).

Perhaps the most powerful analgesic drugs in our armamentarium are the opioids. Unfortunately, very large inter-individual differences in opioid require-

---

Received 8 April, 2010; accepted 30 July, 2010.

Address for correspondence: Martin S. Angst, MD, Department of Anesthesia, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA, USA. E-mail: ang@stanford.edu

ments exist even in patients suffering from similar conditions (Aubrun et al., 2003; Carvalho et al., 2006). Many patients experience opioid-related side-effects, and the rate of death from overdose in the setting of pain management has increased rapidly over recent years (Dunn et al., 2010). Also problematic is the abuse potential of opioids, which has risen in parallel with their use as pain relievers. The yearly economic costs of prescription opioid abuse total several billion dollars in the United States alone (Strassels, 2009). While tools are available to estimate the likelihood of adverse behaviors during opioid treatment, these remain crude and are based primarily on demographic information, behavioral observations and substance abuse history (Butler et al., 2008). Given these facts, it is essential that we move towards a better understanding of pain, analgesic responsiveness, susceptibility to side-effects and the propensity of individuals to abuse prescription opioids.

The twin study paradigm is cost effective, provides power and is applicable to carefully controlled protocols. Genetic and environmental contributions to inter-individual differences in drug responses can be estimated. This is critical for designing future studies aimed at identifying specific factors predicting the appropriateness of opioid therapy for individual patients. Evidence of significant heritability is of particular value for justifying larger-scale and expensive genomic studies. However, there have been few attempts to apply the twin study paradigm to interventional and comprehensive pharmacological studies. Twin studies in the past examined environmental and genetic contributions to pain associated with a clinical disease such as low back pain or fibromyalgia (Battie et al., 2007; Fejer et al., 2006; Hartvigsen et al., 2004; Hartvigsen et al., 2005; Markkula et al., 2009; Michalowicz et al., 2000; Zondervan et al., 2005). Few twin studies have focused on assessing pain sensitivity per se under experimental conditions, while no study has examined analgesic interventions (MacGregor et al., 1997; Nielsen et al., 2008; Norbury et al., 2007).

In this report we present our approach for conducting such a study on opioid pharmacology. Most opioid effects were captured by measuring surrogate outcomes of clinical pain, analgesia, sedation, respiratory depression, and affective reinforcing drug effects. Surrogate outcomes offer the advantage of being well-defined, homogenous, and highly quantifiable if studied under laboratory-type conditions. In the case of significant heritability such surrogate measures are good endophenotype-candidates, that is, measurable components along the pathway between a complex clinical phenotype and a distal genotype (Gottesman & Gould, 2003). Endophenotypes offer the promise for more straightforward and successful genetic analysis. Studies, such as the one described here, are better suited than clinical studies for facilitating our understanding of the genetic architecture governing pain and analgesia. Information on procedures, infrastruc-

ture and logistics, and a synopsis of early observations speaking toward the feasibility and utility of the outlined approach should be of interest to investigators considering similarly comprehensive pharmacological studies in twins.

## Materials and Methods

### Setting

Twins underwent a 1-day study protocol in the Human Pain Laboratory of the Department of Anesthesia at Stanford University School of Medicine. The laboratory offered precise temperature and lighting control as well as sound-isolation to comfort level. The room was equipped with (1) an ergonomic treatment chair that allowed positioning subjects comfortably in supine or sitting position according to the experimental needs (Cloud 9, Living Earth Crafts, Vista, CA), (2) a vital signs monitor (Propaq Model 244, Welch Allyn, Beaverton, OR), (3) wall-mounted oxygen supply and suction, (4) a crash cart with a defibrillator, airway management equipment, emergency drugs, and additional oxygen supply, (5) carts containing disposables for specimen collection and drug administration, and (6) equipment to assess pain and respiratory parameters (described in detail below). Resting periods during the study were standardized; subjects listened to music via headphones, the room light was dimmed, any interaction with the subject was avoided, and activities by study staff causing noise or possibly distracting subjects were prohibited.

Study staff consisted of (1) a research associate conducting all experiments, (2) a registered nurse familiar with critical care performing phlebotomies, administering study drug, and monitoring vital signs, and (3) an anesthesiologist overseeing drug infusions and assuring the safety of participants. Vital signs including heart rate (electrocardiogram), blood pressure, respiratory rate, and hemoglobin oxygen saturation were monitored throughout the study. Participants received supplemental oxygen (2l/min) via nasal cannula during the infusion of study medication.

### Recruitment

Recruitment was overseen by SRI International (Menlo Park, CA), which maintains the Twin Research Registry (TRR) (Swan et al., 2004). Recruitment utilized the TRR and an integrated marketing and communications campaign including radio and online advertisement, and a media and community outreach. The use of a diverse portfolio of media channels accounted for the demographic diversity of the targeted twin audience. Specific components of the campaign were (1) advertisements on two separate San Francisco Bay Area radio stations targeting different age groups, (2) community outreach via flyers posted at local high-traffic venues, and information distributed over the Internet (LinkedIn, SRI's newsletters, and SRI's Facebook page), and (3) media outreach via press releases posted on SRI's newsroom and distributed by the Marketwire press release wire

service, and by actively contacting local reporters and news producers to secure their interest in a story about the study.

#### Enrollment Process

Twins were contacted by an experienced recruiter of SRI International who was well familiar with the protocol, the potential risks and the underlying scientific aims of the study. The recruiter had participated in a mock trial covering all study procedures. Participants' eligibility was established and contact information and screening results were forwarded to the study coordinator at Stanford University School of Medicine who scheduled an appointment, provided initial instructions, and arranged for transportation. Written informed consent was obtained on the day of the study. The study was approved by the Institutional Review Boards of Stanford University School of Medicine and SRI International prior to any recruitment effort.

#### Enrollment Criteria

**Inclusion.** (1) age between 18 and 70, (2) monozygotic or dizygotic twins both consenting to study participation, (3) fluent in English language, (4) willing and able to sign an informed consent form and HIPAA authorization and to comply with study procedures, (5) negative urine pregnancy test on study day (premenopausal women), (6) fasted overnight (clear liquids up to 2 hours before drug infusion).

**Exclusion.** (1) Clinically relevant systemic diseases such as psychiatric, neurological, and dermatological conditions interfering with the collection and interpretation of study data, (2) clinically relevant cardio-respiratory diseases causing at least moderate impairment in daily activities, (3) renal and hepatic diseases with functional impairment, (4) morbid obesity and/or history of significant sleep apnea, (5) history of addiction, (6) allergy to study medication, (7) chronic intake of medications with recognized analgesic/antihyperalgesic activity, (8) intake of over-the-counter analgesics within two days prior to the study, (9) Raynaud's disease, (10) pregnancy, (11) other conditions compromising a participant's safety or the integrity of the study.

#### General Experimental Design

A single occasion, two-staged, randomized, double blind and placebo-controlled study design was implemented. The flow of the study is depicted in Figure 1. During the first stage outcome measures and potential covariates affecting pain perception were assessed before administering study medication. During the second stage drug-induced changes in measured outcomes were assessed during the sequential but randomized administration of saline placebo and the opioid agonist alfentanil. Primary outcomes were the analgesic effects of alfentanil. Analgesia was inferred from the drug-induced reduction of participants' sensitivity to heat and cold pressor pain stimuli. Secondary outcomes were the sensitivity to heat and cold pressor pain before drug

administration and opioid effects including sedation, nausea, respiratory depression, pruritus, and drug liking (an index of addiction potential). Relevant covariates known to potentially confound measures of pain and analgesia included demographic factors, depressed mood, anxiety, sleep, and blood pressure.

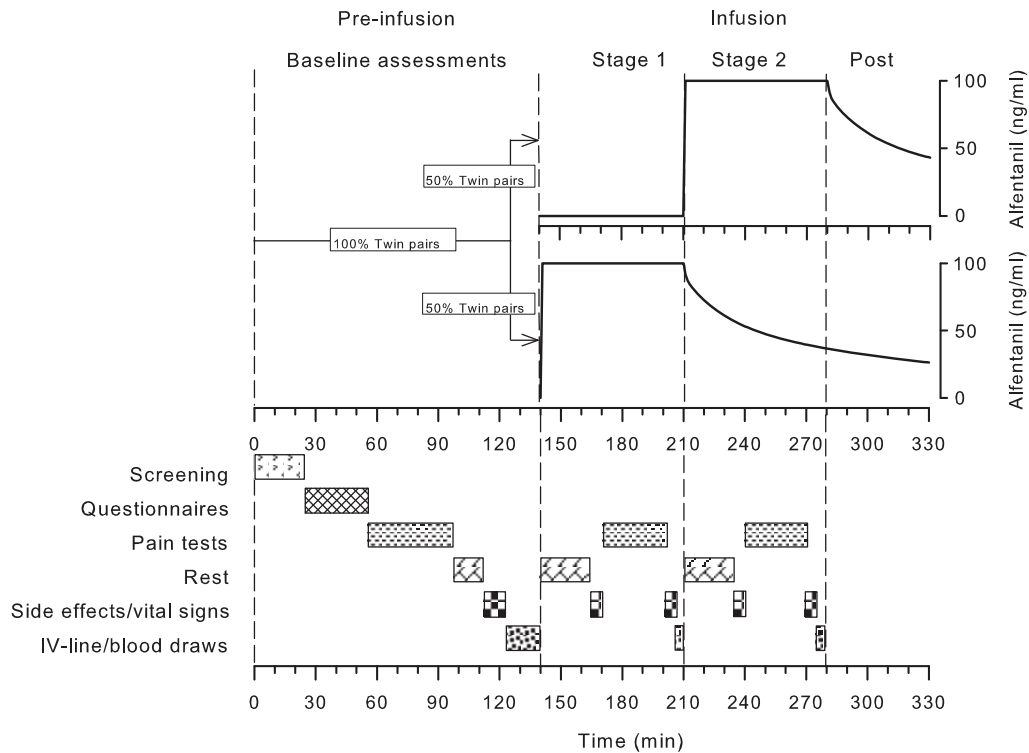
The single-day study design had to control for placebo effects potentially confounding pain and analgesic outcome measures (Figure 1). Traditional designs test for drug and placebo effects on two separate days. However, requiring subjects to return for a second study day was not considered feasible, given the final enrollment target of 125 twin pairs and substantial concerns that such a requirement may significantly hamper our ability to recruit and retain participants. While a single-day study design did not allow assessing placebo effects in all participants, such effects could be assessed in the 50% of twins randomized to receive saline before alfentanil. Randomizing the other 50% of twins to receive alfentanil before saline was necessary to maintain the blinding. However, placebo effects could not be assessed in these twins because residual alfentanil plasma concentrations were still present during the saline infusion. Members of the same twin pair underwent the same infusion sequence.

#### Opioid Administration and Assay

The study was designed to limit the confounding influence of pharmacokinetic variability on outcome measures. A computer controlled infusion (CCI) paradigm was used to quickly achieve and maintain steady-state plasma and effect site concentrations, which allowed performing all test procedures at similar plasma concentrations (Angst et al., 2004). While CCI paradigms keep plasma concentrations stable in individual participants, concentrations still vary among different participants. Therefore, alfentanil plasma concentrations were measured to adjust for inter-individual pharmacokinetic differences in the final analysis.

The  $\mu$ -opioid agonist alfentanil (Janssen Pharmaceutica, Titusville, NJ) was chosen among the class of opioids because of its quick onset of action, a fast recovery from its effects, and a well-validated CCI algorithm for its administration (Angst et al., 2004). Alfentanil was administered intravenously via a computer-controlled infusion pump (Harvard Pump 22, Harvard Apparatus, Inc., South Natick, MA) targeting a steady-state plasma concentration of 100ng/ml. This target concentration produces clinically relevant analgesic effects and significantly attenuates experimental pain without causing harmful side effects (Angst et al., 2004; van den Nieuwenhuyzen et al., 1993). STAN-PUMP using Scott's weight-adjusted pharmacokinetic parameters was the software driving the infusion pump (Scott & Stanski, 1987).

Alfentanil plasma concentrations were assayed at the Clinical Research and Development Unit of the Department of Anesthesia at the University of Colorado Health Sciences Center (Denver, CO). Six



**Figure 1**

Twins were enrolled in a single occasion, double-blinded study paradigm and underwent a target-controlled infusion of the opioid alfentanil and saline placebo in sequential but randomized order. Baseline assessments included covariates affecting pain perception, sensitivity to experimental heat and cold pressor pain, respiratory parameters, mental performance, vital signs, and blood draws. Fifty per cent of twin pairs were allocated to receive alfentanil first and saline placebo second, while the other 50% of twin pairs received alfentanil and saline placebo in reversed order. The alfentanil plasma concentrations for both treatment sequences are depicted in the graph. Assessments during drug infusion included alfentanil-induced analgesia, respiratory depression, sedation, nausea, pruritus, and reinforcing affective effects. The total time requirement for study completion including recovery from drug effects was an estimated 5.5 hr.

milliliters of venous blood were drawn into heparinized glass tubes, centrifuged, and the plasma was frozen and stored at  $-70^{\circ}\text{C}$  until assayed. Using LC/LC-MS/MS the lower limit of quantitation was 1.25 pg/ml with a 1000-fold linear range ( $r = 0.99$ ), and an intra- and between-assay coefficient of variations ranging between 4–16% and 3–14%.

#### Experimental Pain Tests

Experimental pain models offer precise control of the applied noxious stimulus and thereby, can control for the reliable rating of pain by study participants. Experimental pain studies produce data of a quality not achievable by clinical pain studies (Angst et al., 2009; Angst et al., 2001). While experimental and clinical pain are not the same, experimental pain models mimic important mechanistic and phenomenological aspects of clinical pain and produce valid surrogate data (Granot et al., 2003; Werner et al., 2004). Two mechanistically distinct models were used.

**Heat pain:** Heat pain evoked by slowly raising temperatures is mediated by unmyelinated nociceptive fibers and is potently affected by opioids (Yarnitsky et al., 1992). Heat pain was induced with a thermal sensory analyzer (TSA-II, Medoc Advanced Medical

Systems, Durham, North Carolina). A  $3 \times 3$  cm thermode was placed in contact with skin at the volar forearm. Starting at  $35^{\circ}\text{C}$ , the thermode temperature was increased at a rate of  $1^{\circ}\text{C}/\text{s}$ . Study participants pushed a button of a hand-held device at the onset of pain. This procedure was repeated 4 times with an interstimulus interval of 30 s. The average temperature eliciting pain was recorded as the pain threshold.

**Cold pressor pain:** Cold pressor pain is thought to mimic important qualities of clinical pain, since verbal descriptors for both types of pain are strikingly similar (Chen et al., 1989). Cold pressor pain is also mediated by small unmyelinated nociceptive fibers but is more sustained than heat pain and is associated with a much stronger affective response (Rainville et al., 1992). The cold-pressor pain model can be viewed as a tool examining an integrated pain response with a strong affective component, while the heat pain model is better suited to explore sensory-discriminative aspects of pain. Sensitivity to cold pressor pain was tested by having subjects immerse their hand to the wrist in ice-water ( $1\text{--}2^{\circ}\text{C}$ ) continuously recirculated within a 12-liter container. The palm of the hand was in full contact with the bottom of the container. Subjects were asked to indicate the onset of dull,

aching pain typically perceived in the wrist and to withdraw the hand once pain became intolerable. The time to the onset of pain was recorded as the pain threshold and the time to withdrawing the hand was recorded as the cold pressor pain tolerance.

#### Non-Analgesic Opioid Effects

Considering that non-analgesic opioid effects were secondary outcomes, techniques used for their assessment were relatively brief to avoid overburdening study participants with extraneous test procedures.

**Sedation.** Sedative opioid effects were assessed with the trail-making test (TMT) (Angst et al., 2004; Oswald and Roth, 1987). The TMT is a paper-and-pencil test consisting of 4 different matrices listing numbers 1–90 in a 9 × 10 format. Subsequent numbers are located in neighboring rows or columns. Matrices were allocated randomly. Subjects had to connect numbers 1–90 as quickly as possible and the time to completion was recorded.

**Respiratory depression.** Respiratory depression was quantified by measuring changes in partial pressure of transcutaneous carbon dioxide (CO<sub>2</sub>) with aid of a pO<sub>2</sub>/pCO<sub>2</sub>-electrode (Perimed Inc., North Royalton, OH) mounted to the anterior chest wall. While measured tissue pCO<sub>2</sub> is higher than the arterial pCO<sub>2</sub>, the two measures correlate strongly and relative changes match closely (Tobias, 2003).

**Nausea.** Participants rated as to how much they felt nauseated during the infusion on a 100-mm visual analogue scale (VAS) anchored by the words *Not at all* (0) and *As much as possible* (100).

**Pruritus.** Participants rated as to how much they felt itchy during the infusion on a 100-mm VAS anchored by the words *Not at all* (0) and *As much as possible* (100).

**Positive affective response.** Measurements of subjective reinforcing drug effects predict a drug's abuse and addiction potential (Epstein et al., 2006). Among these measurements 'drug liking' has one of the best predictive values (Haertzen et al., 1983). Drug-liking was assessed on a 100-mm VAS anchored by the words *Not at all* (0) and *Like very much* (100).

#### Covariates

Several psychological and physiological factors can affect the perception of pain. Assessing these factors as potential covariates enhances the power for detecting heritability of pain and analgesic outcomes.

**Depression.** Depression is associated with more frequent and severe pain (Lautenbacher et al., 1999). Depressive symptom severity was assessed with the Beck Depression Inventory (BDI), a self-reported questionnaire that takes 5–10 minutes to complete and yields a single score between 0–63 (0–13 = *No depression*, 14–19 = *Mild depression*, 20–28 = *Moderate depression*, > 29 = *Severe depression*) (Beck et al., 1996). The BDI is well validated and widely used in studies of pain.

**Anxiety.** Anxiety correlates with the severity of pain (Taenzer et al., 1986). Anxiety was assessed with the Profile of Mood States (POMS), a self-reported questionnaire that evaluates six dimensions of mood (anxiety, depression, anger, vigor, inertia, and bewilderment) and takes about 3–5 min to complete (McNair et al., 1992). Subjects rated 65 mood-related adjectives on a 5-point scale (0 = *Not at all*, 1 = *A little*, 2 = *Moderately*, 3 = *Quite a bit*, 4 = *Extremely*). The POMS yields a total score and sub-scores for each dimension of mood (anxiety subscore range: 0–36). The POMS is well validated and widely used in studies of pain.

**Sleep.** Disturbed sleep is associated with increased pain sensitivity (Onen et al., 2001; Raymond et al., 2001). Sleep quality during the month preceding the study was assessed with the Pittsburgh Sleep Quality Index (PSQI), a self reported questionnaire that assesses seven components of sleep (quality, latency, duration, efficiency, disturbance, medication, and daytime dysfunction) and takes 3–5 min to complete. The total score ranges between 0–21, and a value > 6 is indicative for sleep disturbance. The PQSI is well validated and widely used (Backhaus et al., 2002; Buysse et al., 1989). Sleep quality and quantity during the night before study participation was assessed on bipolar 100 mm VAS (range -50–50 mm) anchored by the words 'worst/best imaginable' and 'worst/optimal' (Singh et al., 1997).

**Blood pressure.** Blood pressure and pain sensitivity are negatively correlated (Bruehl et al., 1992; Campbell et al., 2003). Blood pressure was measured twice at each arm (Propaq Model 244, Welch Allyn, Beaverton, OR) after a 15-minute resting period in semi-recumbent position. The median was recorded.

**Demographics.** The perception of pain can be affected by several recorded demographic factors including race, ethnicity, gender, age, and educational status (Edwards et al., 2001; Gagliese and Katz, 2003; Taenzer et al., 2000).

**Blood sampling.** A total of 112 ml of blood were collected from each twin to assay for alfentanil plasma concentrations, perform genotyping including final zygosity testing, and perform future DNA, RNA, and proteomic analysis. Preliminary zygosity was assessed with a questionnaire providing 95–99% accuracy (Christiansen et al., 2003; Swan et al., 2004). Samples are stored centrally at the Human Immune Monitoring Core of Stanford University School of Medicine.

**Discharge criteria and follow-up after study completion.** Discharge criteria were those used for patients undergoing non-invasive, ambulatory procedures requiring sedation (e.g., bronchoscopy): (1) Blood pressure ± 20% of baseline, (2) hemoglobin oxygen saturation > 95%, (3) awake, (4) no or mild nausea, (5) no vomiting, (6) able to urinate, and (7) prearranged transport home available. Several weeks after study participation twins were contacted by SRI staff to rate the

severity of possible adverse experiences during the study including phlebotomy, nausea, vomiting, dizziness, and disorientation.

**Data storage.** Data was stored in an Oracle database maintained by the Information Resources and Technology Group at Stanford University School of Medicine. The database is secured physically and electronically, is compliant with the Health Insurance Portability and Accountability Act, and is backed up automatically and housed offsite for disaster recovery.

**Statistical analysis.** Data are presented as mean and standard deviation (SD) or as median and inter-quartile range (IQR). Summary statistics, parametric or non parametric hypothesis testing with paired test procedures, and correlation analysis on continuous and ranked data were performed in Systat Version 13 (Chicago, IL). An alpha level of  $p < .05$  indicated statistical significance.

An interim analysis in 39 monozygotic (MZ) and 22 dizygotic (DZ) pairs (questionnaire-based zygosity testing) provided estimates on familial aggregation and heritability of heat and cold pressor pain sensitivity (Stata Version 11, StataCorp LP, College Station, TX). The difference in the number of subjects used for inferring heritability and reporting demographic and other outcome variables (122 versus 159) is related to the fact that study efforts continued, while the heritability analysis was structured and run in a data set of 122 twins. Since presented heritability results are not final regardless of whether the data set consists of 122 or 159 twins, the analysis was not re-run for all 159 twins. Only previously studied non-pharmacological outcomes (pain sensitivity) were analyzed because consistency with previous results would further support the utility of our study algorithm. Expanding the heritability analysis to include pharmacological phenotypes would be beyond the scope of this manuscript. As importantly, little is known about the heritability of studied pharmacological phenotypes and therefore, first estimates of heritability will be most accurate if made in the complete data set.

A maximum likelihood approach for multi-level, mixed effects models with two random effects was used (family and zygosity nested within family). DZ pairs share the family effect, while MZ pairs share both effects. The mixed effects model is mathematically identical to the traditional ACDE twin model allowing for additive genetic effects (A), common environmental effects shared within families (C), dominance effects at a single locus or other genetic interactions among multiple variants (D), and random effects (E). The family effect in the mixed effects model above encompasses variance component terms C,  $1/2A$  and some portion of D in the traditional ACDE model. The  $1/2A$  arises because DZ twins are on average 50% identical by descent. The exact portion of D depends on unspecified interactions between the alleles at a single locus or among different

loci. The zygosity effect in the mixed effects model encompasses those genetic effects shared by MZ twins that exceed the genetic effects shared by DZ twins ( $1/2$  of variance component of A and the remainder of D).

The twin study design provides information to estimate exactly two correlational parameters, the MZ and the DZ intra-class correlations (ICCs). Thus, the A, C and D variance components in the full ACDE model are not simultaneously identifiable. The ACE, ADE and ACDE models differ in the constraints imposed on the ICCs. All three models assume that the MZ ICC is no less than the DZ ICC. In the ACE model, the MZ ICC is no more than twice the DZ ICC. In the ADE model, the MZ ICC is no less than twice the DZ ICC. The ACDE model has neither constraint. As described below, we use likelihood ratio tests appropriately adjusted for boundary conditions to compare these models. We follow the usual practice estimating narrow-sense (additive) heritability under the ACE model to avoid identifiability problems. The CE and E models were fit with a single family-level and no random effects, respectively.

A likelihood ratio test (2df) compared models ACDE and E to test for familial aggregation. A one-sided test (1df) compared models ACDE and CE to test for heritability, i.e., MZ intra-class correlation larger than DZ intra-class correlation. A one-sided test (1df) separately compared models AE and CE to model E. The D and C variance components were tested simultaneously, using an unadjusted two-sided (1df) test to compare models ACDE and AE.

Heritability and the effect of shared family environment (C) were estimated as non-linear combinations of the estimated variance components from the ACE model with 95% confidence intervals produced by the delta-method as implemented by the 'nlcom' function in Stata. Narrow-sense definitions were used in which heritability =  $V_A/(V_C+V_A+V_E)$ , and shared environmental effects =  $V_C/(V_C+V_A+V_E)$ , where  $V_C$ ,  $V_A$ , and  $V_E$  are the estimated variance components contributed by the shared family environment, additive genetic factors, and random error, respectively. Intra-class correlations were similarly estimated from the full ACDE model. All models were adjusted for age and sex. An alpha level of  $p < .05$  indicated statistical significance.

#### Sample Size

The final enrollment goal is 80 MZ and 45 DZ twin pairs. This was based on a power analysis considering familial aggregation first and heritability second. Power was approximated based on Fisher's log transformation for the Pearson correlation coefficient between measurements for twin A and twin B at an alpha level of  $p < 0.05$ . Eighty-five MZ pairs yielded a power of 0.8 to detect familial aggregation of 30% or more. Eighty-five MZ and 40 DZ pairs yielded a power of 0.5 to detect heritability of 50% or more if common environmental effects account for a modest 20% of the variance. However, these calculations are best approximations because the precise power will

depend on the proportion of the response variance that is due to additive genetic effects (A), shared environmental effects (C), and individual environmental and random variation (E).

## Results

### Subjects

Results are reported for 159 participants. Sixty-three percent were women and 37% were men. The mean (SD) age was  $36 \pm 15$  years. Fifty-two per cent were college graduates suggesting that it may be easier to recruit twins with higher education to interventional and comprehensive pharmacogenomic studies. Sixty-five per cent of the twin pairs were monozygotic, 32% were dizygotic, and 3% were of undetermined zygosity based on a questionnaire providing 95–99% accuracy (Christiansen et al., 2003; Swan et al., 2004). Demographic details are depicted in Table 1.

### Recruitment Efforts

From September 2008 until December 2009, 591 individuals were contacted for study participation. Of these, 179 were enrolled (30%) and 159 (27%) had completed the study by December 2009. Thirty-one per cent of contacted twins were not eligible (8% health-related, 5% medication-related, 17% co-twin ineligible, 1% other reasons) and 39% declined participation (16% co-twin, spouse, or parent refused, 12% living too far, 4% test medication or cannulation of vein, 7% other reasons).

Twins were recruited through the following approaches: (1) 37% members of TRR, (2) 33% by radio advertisement, (3) 12% by the printed press, (4)

11% by other twin study participants, (5) 4% by television news segments, and (6) 3% by Internet advertisement or locally distributed posters. Radio advertisements were 30-second segments broadcast 84 times for 5 weeks during winter 2009 on a station targeting the middle-aged population, and broadcast 176 times for 4 weeks during spring 2009 on a station targeting younger adults. Articles about the twin study were published in five local newspapers. Two 3-minute news segments were produced and shown several times by two local television stations.

The median (IQR) time between recruitment and study participation was 61 (31–85) days. The median distance between participants' home and the study center was 36 (25–60) miles. The median cost for transportation was \$60 (44–100). Participants were either transported by friends or family members who were reimbursed at \$0.5/mile, or by taxi. Sixty-four per cent of the study sessions took place on weekdays, while 36% took place during weekends.

### Study Conduct

Time requirements to complete the study are depicted in Table 2. The median duration was about 5 hr. However, 8% of participants required more than 6 hr and up to 7.5 hr. The two major reasons for significantly prolonged studies were difficulties establishing intravenous access and delayed recovery from drug effects.

**Intravenous access.** The most extreme examples were five twins who required between 80–130 min (including resting periods). Such efforts seemed justified given that twins constitute a sparse study population. Initial strategies used to facilitate intravenous access were (1)

**Table 1**

Participant Characteristics

Category	Subcategories	%
Race	American Indian	4
	Asian	15
	Pacific Islander	1
	African American	4
	Caucasian	76
Ethnicity	Hispanic	17
	Non-Hispanic	83
Marital status	Single	59
	Married	38
	Divorced	2
	Widowed	1
Education	Some high school	0
	High school graduate	9
	Some college	39
	College graduate	52
Profession	Working	69
	Student	20
	Retired	4
	Unemployed	7

**Table 2**

Time Requirements

Procedure	Time (min)		
	Median	25–75%	5–95%
Total study <sup>1</sup>	305	285–325	257–372
Pre-infusion phase	151	138–164	124–200
Screening <sup>2</sup>	22	19–25	13–32
Questionnaires <sup>3</sup>	18	15–21	11–30
Pain tests <sup>4</sup>	35	30–41	23–50
IV & blood sampling <sup>5</sup>	25	16–35	11–59
Placebo infusion	56	49–65	43–77
Pain tests	20	17–25	15–30
Alfentanil infusion	61	55–69	46–82
Pain tests	22	18–27	16–33
Post infusion phase	28	19–43	12–73

Note: <sup>1</sup> Entry of study site until discharge including pre-infusion phase, infusion phase and recovery phase

<sup>2</sup> Confirm eligibility, answer questions, written informed consent, demographics, medical history, reimbursement information

<sup>3</sup> Profile of Moods State, Beck Depression Inventory, Pittsburgh Sleep Quality Index

<sup>4</sup> Including training in heat and cold pressor pain tests before collecting pain data

<sup>5</sup> Insertion of an 18 or 20G venous cannula

instructing twins to drink clear fluid in the morning to avoid significant dehydration, (2) warming of the forearm with blankets, and (3) cutaneous administration of nitroglycerin ointment at the insertion site. Despite such efforts line placement remained difficult in some participants, which triggered the use of a non-invasive ultrasound-guided cannulation technique (Site Rite IV, Bard Access Systems, Salt Lake City, UT). This greatly shortened time requirements and prevented any failed placement.

**Delayed recovery:** Four twins had significantly prolonged recoveries ranging between 90–120 min.

A critical element of studies collecting subjective pain data is adequate training of participants to ascertain consistent performance. Participants had to be successfully trained on the day of the study without unduly prolonging study duration. A consistent performance was defined as testing within 0.7°C for the heat pain threshold and within 20% for the cold pain tolerance on two consecutive test sessions. All participants except one were successfully trained within 60 min. For the heat pain testing, 2 cycles were required for 47%, while 3, 4, and 5 cycles were required for 48%, 8%, and 1 % of participants. For the cold pain testing, 2 cycles were required for 61%, while 3 and 4 cycles were required for 34%, and 5% of the participants, respectively.

#### Placebo Effects

Placebo effects on analgesic outcome measures were assessed in subjects randomized to receive saline first and alfentanil second during the infusion phase. The median (IQR) heat pain threshold was 48.6°C (47.3–49.1) at baseline and 48.5°C (47.4–49.5) during saline placebo infusion. The median cold pressor pain threshold and pain tolerance were 8s (5–12) and 16s (11–23) at baseline and 7s (5–12) and 14s (8–23) during saline placebo infusion. Administration of saline placebo did not produce significant analgesic effects.

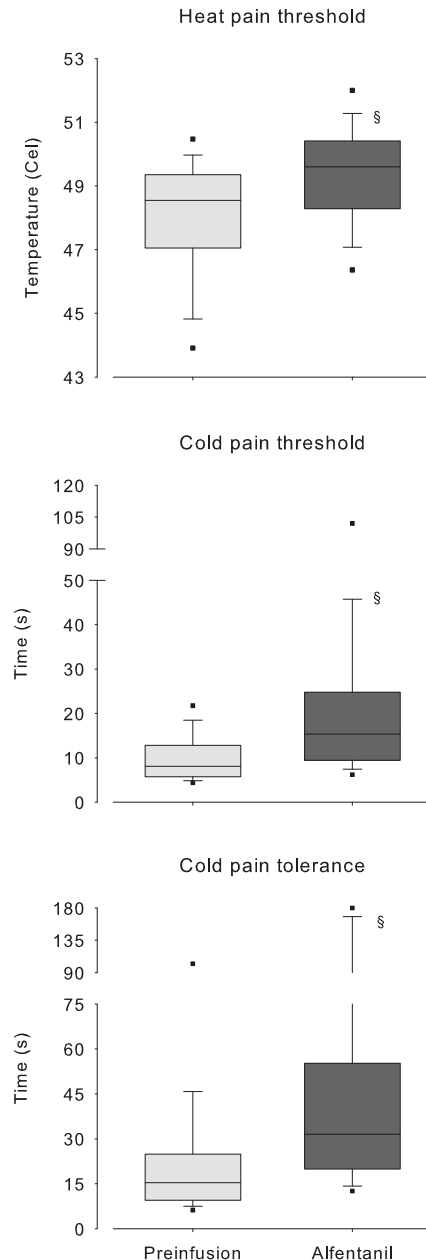
Placebo effects on measures of sedation and respiratory depression are discussed in the section presenting results for non-analgesic opioid effects. Analgesic and non-analgesic opioid effects measured during placebo administration were all significantly different from the effects measured during alfentanil administration ( $p < .001$ ).

#### Alfentanil Plasma Concentrations

Plasma concentrations were determined in 128 twins in an interim analysis. The median (IQR) concentration was 62ng/ml (51–78), which is 38% lower than predicted. This deviation is larger than the 20% previously reported (Angst et al., 2004). The deviation correlated significantly with the body mass index (BMI), that is, a lower BMI was associated with a larger deviation ( $R = 0.55$ ,  $p < .001$ ). Nevertheless, actual plasma concentrations produced sizable drug effects. The median difference within twin pairs was 13% (6–25).

#### Analgesic Opioid Effects

Infusion of alfentanil caused significant analgesia ( $p < .001$ ) to heat and cold pressor pain (Figure 2). The median (IQR) heat pain threshold increased from 48.6°C (47.1–49.3) to 49.6°C (48.3–50.4). The median



**Figure 2**

Infusion of alfentanil caused significant analgesia to experimental heat and cold pressor pain ( $\$$ ;  $p < .001$ ). The heat pain threshold was defined as the minimum temperature of a thermode in contact with skin causing the onset of pain. The cold pressor pain threshold was defined as the time to the onset of pain after immersion of the hand in an ice-water container. The cold pressor pain tolerance was defined as the time to withdrawing the hand from the ice-water container because of intolerable pain. Results are depicted as box plots. The line within the box marks the median, the boundaries of the box indicate the 25th and 75th percentiles, the error bars indicate the 10th and 90th percentiles, and the dots indicate outliers.



cold pressor pain threshold increased from 8s (6–13) to 15s (9–25), and the median cold pressor pain tolerance increased from 17s (11–26) to 32s (20–55).

**Non-Analgesic Opioid Effects**

**Sedation.** Alfentanil significantly ( $p < .001$ ) increased the median time (IQR) to complete the TMT from 63s (56–75) to 70s (59–87) (Figure 3). The time to complete the trail making test at baseline and during saline placebo administration in subjects randomized to receive saline first and alfentanil second during the infusion phase was 67 s (56–78) and 67 s (58–79) indicating absence of a placebo or learning effect.

**Respiratory depression.** Alfentanil significantly ( $p < .001$ ) increased the median transcutaneous partial pressure of carbon dioxide (tpCO<sub>2</sub>) from 40.3mmHg (37.6–42.9) to 47.4mmHg (43.4–52.0) (Figure 3). The tpCO<sub>2</sub> at baseline and during saline placebo administration was 39.8mmHg (37.6–42.4) and 38.6mmHg (37.2–41.2) in subjects randomized to receive saline first and alfentanil second during the infusion phase indicating absence of a placebo effect.

**Nausea.** Alfentanil caused nausea in 54% of participants. The median VAS score was 30 (15–60). Only 3% of participants reported nausea during placebo administration (VAS 2; 2–4).

**Pruritus.** Alfentanil caused pruritus in 67% of participants. The median VAS score was 30 (10–50). Only 16% of participants reported pruritus during placebo administration (VAS 2; 2–26).

Positive affective response: Alfentanil caused ‘drug-liking’ in 81% of participants. The median VAS score of maximum liking was 80 (60–90). Only 16% of participants reported drug liking during placebo administration (VAS 63; 40–75).

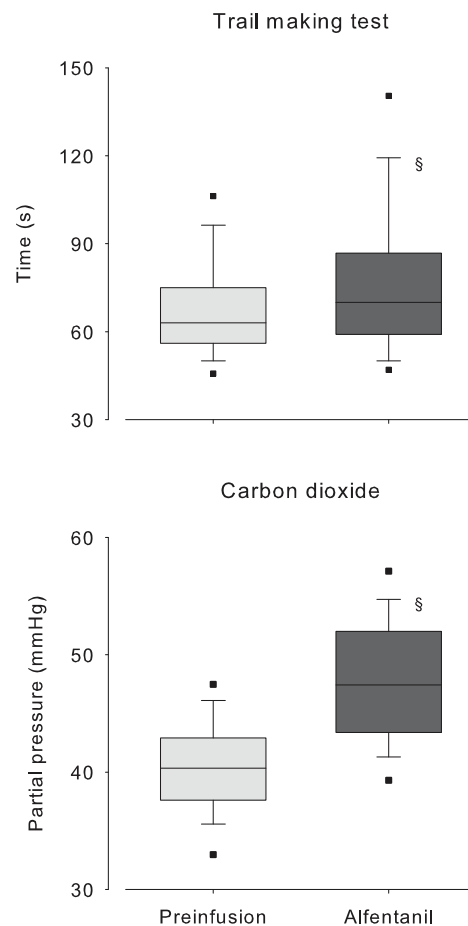
**Covariates**

Results are summarized in Table 3. A preliminary analysis focused on detecting significant correlations between covariates and measures of pain sensitivity.

**Depression.** Only seven participants had a BDI score compatible with mild ( $n = 5$ ) or moderate/severe ( $n = 2$ ) depression. Given the low prevalence, depression unlikely confounds outcome measures.

**Anxiety.** Participants’ scores were low. The highest score recorded in this study closely matched the mean score reported in the normative database (McNair et al., 1992). Lower than normal anxiety scores may reflect a recruiting bias inherent to studies using invasive procedures. Anxiety unlikely confounds outcome measures.

**Sleep.** About 30% of participants had an abnormal PQSI > 6 (range 7–13) and provided a negative VAS score for sleep quality and/or quantity; both suggestive for impaired sleep. No significant correlations were detected between PQSI/VAS and measures of pain sensitivity. However, given the prevalence of an abnormal PQSI, sleep disturbance may be a relevant covariate in the final analysis considering additional outcomes including analgesic opioid effects.



**Figure 3**

Infusion of alfentanil caused significant sedation (§;  $p < .001$ ) as indicated by an increased time to complete the trail-making test (TMT). The TMT is a paper-and-pencil test asking subjects to connect numbers 1–90, assorted semi-randomly, as quickly as possible and the time to completion is recorded. Similarly, alfentanil caused significant respiratory depression (§;  $p < .001$ ) as indicated by an increase in carbon dioxide (CO<sub>2</sub>). The partial pressure of CO<sub>2</sub> was measured with aid of a pO<sub>2</sub>/pCO<sub>2</sub>-electrode mounted to the anterior chest. Results are depicted as box plots. The line within the box marks the median, the boundaries of the box indicate the 25th and 75th percentiles, the error bars indicate the 10th and 90th percentiles, and the dots indicate outliers.

**Table 3**

Covariate	Median	25–75%	5–95%
<b>Sleep</b>			
PSQI (0–17)	5	3–7	2–10
VAS quality (-50–50)	12	-4–27	-26–42
VAS quantity (-50–50)	8	-1–26	-24–44
<b>Depression</b>			
BDI (0–63)	3	1–7	0–14
<b>Anxiety</b>			
POMS (0–36)	4	2–7	0–11
<b>Blood pressure</b>			
Systolic (mmHg)	116	108–124	98–141
Diastolic (mmHg)	65	58–72	51–83

**Blood pressure.** Blood pressure varied widely in participants. A modest but significant negative correlation was detected between diastolic blood pressure and heat pain sensitivity ( $R = 0.26$ ;  $p < .001$ ) but not cold pressor pain sensitivity. Blood pressure likely is a relevant covariate in the final analysis.

**Demographics.** A significant negative correlation was detected between age and heat pain sensitivity ( $R = 0.41$ ;  $p < .001$ ) but not cold pressor pain sensitivity. No significant correlations were detected between pain sensitivity and gender, race, ethnicity, or education.

#### Safety Parameters

Safety parameters are summarized in Table 4. The respiratory rate significantly decreased during the infusion of alfentanil ( $p < .001$ ). Prophylactic administration of oxygen via nasal cannula (2–3l/min) sufficed to maintain adequate hemoglobin oxygen saturation. Blood pressure and heart rate remained stable throughout the different infusion phases.

#### Administration of Rescue Medication

Asymptomatic episodes of bradycardia triggered the intravenous administration of 0.2mg glycopyrrolate on three occasions. Anti-emetic therapy occurred in 63% of the participants with 57% requiring one or two intravenous doses of 4mg ondansetron, and 6% requiring an additional intravenous dose of 10mg metoclopramide or application of a 1.5 mg scopolamine patch.

#### Follow-Up

So far, information has been collected in 79 participants (Table 5). It is remarkable that 96% of these twins indicated that they would recommend study participations to others when considering that a signif-

icant number experienced opioid-mediated side effects and/or some stress related to the insertion of an intravenous cannula.

#### Familial Aggregation and Heritability of Pain Sensitivity

The threshold to cold pressor pain showed significant familial aggregation ( $p < .001$ ) and heritability ( $p = .015$ ). Intraclass correlations were 0.24 and 0.64 for DZ and MZ pairs under the ACDE model adjusted for age and sex. The ACDE and AE models were not significantly different ( $p = .65$ ), consistent with the absence of a family environmental effect (C) and a genetic interaction effect (D). Accordingly, the AE model provided the best fit. The variance component attributable to heritability was 63% (44–82%; 95%-confidence interval) in an ACE model ( $C = 0\%$ ). Similarly, the tolerance to cold pressor pain showed significant familial aggregation and heritability ( $p < .001$ ). Intraclass correlations were 0.15 and 0.76 for DZ and MZ pairs. The ACDE and AE models were not significantly different ( $p = .17$ ) and the AE model provided the best fit. The variance component of the cold pressor pain tolerance attributable to heritability was 75% (61%–89%; 95%-confidence interval) under the ACE model ( $C = 0\%$ ). The variables age or sex had no significant effects on cold pressor pain sensitivity.

In contrast, the heat pain threshold did not show significant familial aggregation ( $p = .11$ ), or heritability ( $p = .35$ ) when comparing the ACDE with the E model. Intra-class correlations for the heat pain threshold were a modest 0.22 and 0.31 for DZ and MZ pairs. However, tests comparing either the CE model or the additive genetic AE model to the E model both revealed significant results ( $p = .019$ ). The

**Table 4**

Vital Signs

Parameter <sup>1</sup>	Pre-infusion	Placebo	Alfentanil
Respiratory rate <sup>2</sup> (min <sup>-1</sup> )	15 (14–18)	14 (12–16)	11 (9–13)
Oxygen saturation (%)	98 (97–100)	100 (100–100)	100 (99–100)
Heart rate (min <sup>-1</sup> )	59 (53–66)	56 (51–62)	60 (55–68)
Blood pressure			
Systolic (mmHg)	116 (108–124)	115 (107–124)	118 (107–130)
Diastolic (mmHg)	65 (58–72)	63 (54–71)	62 (54–69)

Note: <sup>1</sup> Mean and interquartile range

<sup>2</sup> Pre-infusion > alfentanil ( $p < .001$ ).

**Table 5**

Retrospective Rating of Adverse Experiences (%)

Parameter	None	Mild	Moderate	Severe
Phlebotomy	70	7	9	14
Nausea	32	20	21	27
Vomiting	80	4	2	13
Dizziness	49	21	15	15
Disorientation	76	14	5	5

variance component of the heat pain threshold was an estimated 18% for heritability and 12% for familial aggregation under the ACE model. However, the 95% confidence interval of both estimates included zero. The final analysis of the full data set should provide a more definite answer regarding a possible genetic contribution to heat pain response variance. Environmental and dominance effects were not significant ( $p = .73$ ). Age significantly correlated with an increase in heat pain threshold under the ACDE model ( $p < .001$ ), while no significant association was detected for gender.

## Discussion

The use of opioids for improved pain control has increased dramatically over the past decade (Dunn et al., 2010). However, this change in practice is not without its drawbacks. Individual responses to opioids vary more than tenfold, opioid therapy is associated with cumbersome and sometimes serious side-effects, abuse of prescription opioids has become more prevalent, and the number of deaths from opioid overdose has increased sharply (Aubrun et al., 2003; Dunn et al., 2010; Smith et al., 2009). There is widespread hope that studies characterizing the genetic and environmental factors contributing to analgesic responsiveness, the propensity to experience side-effects and the likelihood of developing addiction will provide the basis for appropriate patient selection. In this respect, twin studies are uniquely suited to estimate environmental and genetic contributions to inter-individual differences in opioid responses. However, applying the twin algorithm to complex and potentially harmful studies comprehensively characterizing pharmacological response profiles is without much precedent. Methods and results reported here highlight the feasibility of such studies and emphasize aspects relevant to their successful completion.

The feasibility of this study depended on several factors. A key concern was the recruitment of a sufficient number of twins. However, about 30% of all contacted twins completed the study. A likely reason for this success rate was an experienced and committed recruiter who had undergone all test procedures in a mock trial. Subjects frequently remarked on the accuracy of information provided by this team member. Another likely reason was the orchestrated media and advertising campaign bringing more twins to the study than contacting twins already registered in TRR. Unexpectedly, 11% of twins were referred by twins who had already completed the study. This is consistent with 96% of twins indicating that they would recommend participation in the study to others, despite suffering from side-effects including nausea, vomiting, and dizziness. Our experience highlights that an invasive and intensive protocol doesn't necessarily deter twins from study participation. In fact, many twins commented on their genuine interest in helping resolve an important scientific question and

advance health care. In this respect, moderate travel distances (20–60 miles) were also not barriers to study participation. Finally, the availability of weekend sessions is important because a third of our studies were conducted on Saturdays or Sundays.

The design of a protocol allowing collection of comprehensive data sets without overburdening participants and risking their early withdrawal was a challenge. At this point we have completed 100% of all initiated study sessions without the occurrence of serious adverse events necessitating early study termination. In part, this record can be attributed to the selection of an appropriate opioid dose (Angst et al., 2004). However, the assembly of staff with credentials required for administering opioid infusions in a clinical context was also a likely key factor. Vigilant monitoring of vital signs, prompt interventions at the occurrence of adverse events, and appropriate reassurance of participants by staff nurses and physicians helped retain study participants. This is not to say that challenges were not faced. For example, establishing venous access in several twins was quite difficult despite the involvement of experienced nurses and anesthesiologists. This highlights the need for proficient study personnel with pertinent experience. Some twins experienced a prolonged recovery of about two hours after the opioid infusion. Studies administering drugs with a significant side effect profile need infrastructure and personnel to accommodate participants for flexible amounts of time. Time requirements for procedures other than line placement and recovery from drug effects were much more predictable.

Based on published clinical reports examining the analgesic efficacy of alfentanil, opioid effects were studied at a target plasma concentration of 100ng/ml (Angst et al., 2004; van den Nieuwenhuyzen et al., 1993). Considering previous work, we expected actual alfentanil plasma concentrations to be about 20% lower than target concentrations (Angst et al., 2004). However, actual plasma concentrations measured in this study of more than one hundred subjects were 38% lower than target concentrations. This finding corroborates the fact that target controlled delivery of alfentanil with aid of STANPUMP and Scott's weight-adjusted pharmacokinetic parameters consistently produces plasma concentrations that are lower than predicted (Scott & Stanski, 1987). Investigators using this algorithm should anticipate up to a 40% discrepancy between actual and target plasma concentrations in subjects with a normal BMI. Despite lower than predicted alfentanil plasma concentrations, robust analgesic responses to heat and cold pressor pain were measured. Similarly, significant reinforcing and aversive opioid responses were evoked by achieved plasma concentrations. These findings instill confidence with respect to the final analysis estimating genetic and environmental contributions to opioid response variance.

Twin studies in the past examined environmental and genetic contributions to pain associated with a

clinical disease including low back pain, neck pain, pelvic pain, fibromyalgia, and temporomandibular joint dysfunction (Battie et al., 2007; Fejer et al., 2006; Hartvigsen et al., 2004; Hartvigsen et al., 2005; Markkula et al., 2009; Michalowicz et al., 2000; Zondervan et al., 2005). Relatively few twin studies have focused on assessing pain sensitivity per se under experimental conditions, while no study has examined analgesic interventions (MacGregor et al., 1997; Nielsen et al., 2008; Norbury et al., 2007). However, such studies generally requiring a laboratory-type setting seem better suited than clinical studies to facilitate our understanding of fundamental pain and analgesic mechanisms. Previous nonpharmacological studies in twins examining sensitivity to cold pressor pain reported that 54% of the variance was due to genetic effects (Nielsen et al., 2008). This is similar to our results suggesting that 63% and 75% of the variance to cold pressor pain threshold and tolerance was due to genetic effects. Two other studies determined the genetic contribution to heat pain sensitivity. Our estimate of 18% is similar to the 25% reported by Nielsen but significantly lower than the 53% reported by Norbury (Nielsen et al., 2008; Norbury et al., 2007). The study by Norbury involved only women possibly explaining the discrepant findings. Taken together, these results suggest that cold pressor pain but not heat pain sensitivity is a good candidate for endophenotype status and support the pursuit of studies examining the molecular genetics and proteomic architecture of cold pressor pain sensitivity.

In summary, we have described methods assisting the planning and execution of complex pharmacological and interventional studies in twins. Documenting feasibility and reporting tangible results in a mid-sized twin study should encourage other investigators considering similar approaches. As we complete enrollment, data regarding the genetic contribution to analgesic, positive affective and aversive opioid effects will become available to guide the design of more detailed pharmacogenomic studies.

### Acknowledgment

This work was supported by the National Institute on Drug Abuse, grant 5R01DA023063-02. The authors thank Jill Rubin (telephone screening and recruitment), Mary McElroy and Lisa Jack (recruitment coordination), Ruth Krasnow (participant data-basing and tracking), Dina Basin (media liaison), and Lucia Panini (community outreach).

### References

- Angst, M. S., Chu, L. F., Tingle, M. S., Shafer, S. L., Clark, J. D., & Drover, D. R. (2009). No evidence for the development of acute tolerance to analgesic, respiratory depressant and sedative opioid effects in humans. *Pain, 142*, 17–26.
- Angst, M. S., Drover, D. R., Lotsch, J., Ramaswamy, B., Naidu, S., Wada, D. R., & Stanski, D. R. (2001). Pharmacodynamics of orally administered sustained-release Hydromorphone in humans. *Anesthesiology, 94*, 63–73.
- Angst, M. S., Ramaswamy, B., Davis, M. F., & Maze, M. (2004). Comparative analgesic and mental effects of increasing plasma concentrations of dexmedetomidine and alfentanil in humans. *Anesthesiology, 101*, 744–752.
- Aubrun, F., Langeron, O., Quesnel, C., Coriat, P., & Riou, B. (2003). Relationships between measurement of pain using visual analog score and morphine requirements during postoperative intravenous morphine titration. *Anesthesiology, 98*, 1415–1421.
- Backhaus, J., Junghanns, K., Broocks, A., Riemann, D., & Hohagen, F. (2002). Test–retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. *Journal of Psychosomatic Research, 53*, 737–740.
- Battie, M.C., Videman, T., Levalahti, E., Gill, K., Kaprio, J. (2007). Heritability of low back pain and the role of disc degeneration. *Pain, 131*, 272–280.
- Beck, A., Steer, R., & Brown, G. (1996). *Beck Depression Inventory — Manual*. San Antonio: The Psychological Corporation.
- Bouhassira, D., Lanteri-Minet, M., Attal, N., Laurent, B., & Touboul, C. (2008). Prevalence of chronic pain with neuropathic characteristics in the general population. *Pain, 136*, 380–387.
- Bruhl, S., Carlson, C. R., & McCubbin, J. A. (1992). The relationship between pain sensitivity and blood pressure in normotensives. *Pain, 48*, 463–467.
- Butler, S. F., Fernandez, K., Benoit, C., Budman, S. H., & Jamison, R. N. (2008). Validation of the revised Screener and Opioid Assessment for Patients with Pain (SOAPP-R). *Journal of Pain, 9*, 360–372.
- Byusse, D. J., Reynolds, C.F., 3rd, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*, 193–213.
- Campbell, T.S., Ditto, B., Seguin, J.R., Sinray, S., Tremblay, R.E. (2003). Adolescent pain sensitivity is associated with cardiac autonomic function and blood pressure over 8 years. *Hypertension, 41*, 1228–1233.
- Carvalho, B., Chu, L., Fuller, A., Cohen, S. E., & Riley, E. T. (2006). Valdecoxib for postoperative pain management after cesarean delivery: A randomized, double-blind, placebo-controlled study. *Anesthesia and Analgesia, 103*, 664–670.
- Chen, A. C., Dworkin, S. F., Haug, J., & Gehrig, J. (1989). Human pain responsiveness in a tonic pain model: psychological determinants. *Pain, 37*, 143–160.
- Christiansen, L., Frederiksen, H., Schousboe, K., Skytthe, A., von Wurmb-Schwark, N., Christensen, K., Kyvik, K. (2003). Age- and sex-differences in the validity of questionnaire-based zygosity in twins. *Twin Research, 6*, 275–278.

- Clark, J. D. (2002). Chronic pain prevalence and analgesic prescribing in a general medical population. *Journal of Pain and Symptom Management*, 23, 131–137.
- Clark, J. D. (2008). The pitfalls of profoundly effective analgesic therapies. *The Clinical Journal of Pain*, 24, 825–831.
- Dolin, S. J., Cashman, J. N., & Bland, J. M. (2002). Effectiveness of acute postoperative pain management: I. Evidence from published data. *British Journal of Anaesthesia*, 89, 409–423.
- Dunn, K. M., Saunders, K. W., Rutter, C. M., Banta-Green, C. J., Merrill, J. O., Sullivan, M. D., Weisner, C. M., Silverberg, M. J., Campbell, C. I., Psaty, B. M., & Von Korff, M. (2010). Opioid prescriptions for chronic pain and overdose: A cohort study. *Annals of Internal Medicine*, 152, 85–92.
- Edwards, R. R., Doleys, D. M., Fillingim, R. B., & Lowery, D. (2001). Ethnic differences in pain tolerance: clinical implications in a chronic pain population. *Psychosomatic Medicine*, 63, 316–323.
- Epstein, D. H., Preston, K. L., & Jasinski, D. R. (2006). Abuse liability, behavioral pharmacology, and physical-dependence potential of opioids in humans and laboratory animals: Lessons from tramadol. *Biological Psychology*, 73, 90–99.
- Fejer, R., Hartvigsen, J., & Kyvik, K. O. (2006). Heritability of neck pain: A population-based study of 33,794 Danish twins. *Rheumatology (Oxford, England)*, 45, 589–594.
- Gagliese, L., & Katz, J. (2003). Age differences in postoperative pain are scale dependent: A comparison of measures of pain intensity and quality in younger and older surgical patients. *Pain*, 103, 11–20.
- Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *The American Journal of Psychiatry*, 160, 636–645.
- Granot, M., Lowenstein, L., Yarnitsky, D., Tamir, A., & Zimmer, E. Z. (2003). Postcesarean section pain prediction by preoperative experimental pain assessment. *Anesthesiology*, 98, 1422–1426.
- Haertzen, C. A., Kocher, T. R., & Miyasato, K. (1983). Reinforcements from the first drug experience can predict later drug habits and/or addiction: Results with coffee, cigarettes, alcohol, barbiturates, minor and major tranquilizers, stimulants, marijuana, hallucinogens, heroin, opiates and cocaine. *Drug and Alcohol Dependence*, 11, 147–165.
- Hartvigsen, J., Christensen, K., Frederiksen, H., & Petersen, H. C. (2004). Genetic and environmental contributions to back pain in old age: A study of 2,108 Danish twins aged 70 and older. *Spine*, 29, 897–901.
- Hartvigsen, J., Petersen, H. C., Frederiksen, H., Christensen, K. (2005). Small effect of genetic factors on neck pain in old age: A study of 2,108 Danish twins 70 years of age and older. *Spine*, 30, 206–208.
- Heiberg, M. S., Nordvag, B. Y., Mikkelsen, K., Rodevand, E., Kaufmann, C., Mowinckel, P., & Kvien, T. K. (2005). The comparative effectiveness of tumor necrosis factor-blocking agents in patients with rheumatoid arthritis and patients with ankylosing spondylitis: A six-month, longitudinal, observational, multicenter study. *Arthritis and Rheumatism*, 52, 2506–2512.
- Heiberg, T., & Kvien, T. K. (2002). Preferences for improved health examined in 1,024 patients with rheumatoid arthritis: Pain has highest priority. *Arthritis and Rheumatism*, 47, 391–397.
- Lautenbacher, S., Sternal, J., Schreiber, W., & Krieg, J. C. (1999). Relationship between clinical pain complaints and pain sensitivity in patients with depression and panic disorder. *Psychosomatic Medicine*, 61, 822–827.
- MacGregor, A. J., Griffiths, G. O., Baker, J., & Spector, T. D. (1997). Determinants of pressure pain threshold in adult twins: Evidence that shared environmental influences predominate. *Pain*, 73, 253–257.
- Markkula, R., Jarvinen, P., Leino-Arjas, P., Koskenvuo, M., Kalso, E., & Kaprio, J. (2009). Clustering of symptoms associated with fibromyalgia in a Finnish Twin Cohort. *European Journal of Pain (London, England)*, 13, 744–750.
- McNair, D. M., Lorr, M., & Droppelmann, L. F. (1992). Profile of mood states. *Educational and Industrial Testing Service*, San Diego, CA.
- Michalowicz, B. S., Pihlstrom, B. L., Hodges, J. S., & Bouchard, T. J., Jr. (2000). No heritability of temporomandibular joint signs and symptoms. *Journal of Dental Research*, 79, 1573–1578.
- Michel, M. Z., Sanders, M. K. (2003). Effectiveness of acute postoperative pain management. *British Journal of Anaesthesia*, 91, 448–44.
- Nielsen, C. S., Stubhaug, A., Price, D. D., Vassend, O., Czajkowski, N., & Harris, J. R. (2008). Individual differences in pain sensitivity: Genetic and environmental contributions. *Pain*, 136, 21–29.
- Norbury, T. A., MacGregor, A. J., Urwin, J., Spector, T. D., & McMahon, S. B. (2007). Heritability of responses to painful stimuli in women: A classical twin study. *Brain*, 130, 3041–3049.
- Onen, S. H., Alloui, A., Gross, A., Eschallier, A., & Dubray, C. (2001). The effects of total sleep deprivation, selective sleep interruption and sleep recovery on pain tolerance thresholds in healthy subjects. *Journal of Sleep Research*, 10, 35–42.
- Oswald, W.D., & Roth, E. (1987). Der Zahlen-Verbindungs-Test (ZVT): Handanweisung. *Hogrefe Verlag fuer Psychologie*, Goettingen.
- Rainville, P., Feine, J. S., Bushnell, M. C., & Duncan, G. H. (1992). A psychophysical comparison of sensory and affective responses to four modalities of experimental pain. *Somatosensory & Motor Research*, 9, 265–277.
- Raymond, I., Nielsen, T. A., Lavigne, G., Manzini, C., & Choiniere, M. (2001). Quality of sleep and its daily

- relationship to pain intensity in hospitalized adult burn patients. *Pain*, 92, 381–388.
- Scott, J. C., & Stanski, D. R. (1987). Decreased fentanyl and alfentanil dose requirements with age. A simultaneous pharmacokinetic and pharmacodynamic evaluation. *Journal of Pharmacology and Experimental Therapies*, 240, 159–166.
- Singh, N. A., Clements, K. M., & Fiatarone, M. A. (1997). A randomized controlled trial of the effect of exercise on sleep. *Sleep*, 20, 95–101.
- Smith, H. S., Kirsh, K. L., & Passik, S. D. (2009). Chronic opioid therapy issues associated with opioid abuse potential. *Journal of Opioid Management*, 5, 287–300.
- Stewart, W. F., Ricci, J. A., Chee, E., Morganstein, D., & Lipton, R. (2003). Lost productive time and cost due to common pain conditions in the US workforce. *Jama*, 290, 2443–2454.
- Strassels, S. A. (2009). Economic burden of prescription opioid misuse and abuse. *Journal of Managed Care Pharmacy*, 15, 556–562.
- Swan, G. E., Benowitz, N. L., Jacob, P., 3rd, Lessov, C. N., Tyndale, R. F., Wilhelmsen, K., Krasnow, R. E., McElroy, M. R., Moore, S. E., & Wambach, M. (2004). Pharmacogenetics of nicotine metabolism in twins: methods and procedures. *Twin Research*, 7, 435–448.
- Taenzer, A. H., Clark, C., & Curry, C. S. (2000). Gender affects report of pain and function after arthroscopic anterior cruciate ligament reconstruction. *Anesthesiology*, 93, 670–675.
- Taenzer, P., Melzack, R., & Jeans, M. E. (1986). Influence of psychological factors on postoperative pain, mood and analgesic requirements. *Pain*, 24, 331–342.
- Tobias, J. D. (2003). Noninvasive carbon dioxide monitoring during one-lung ventilation: end-tidal versus transcutaneous techniques. *Journal of Cardiothoracic and Vascular Anesthesia*, 17, 306–308.
- van den Nieuwenhuyzen, M. C., Engbers, F. H., Burm, A. G., Lemmens, H. J., Vletter, A. A., van Kleef, J. W., & Bovill, J. G. (1993). Computer-controlled infusion of alfentanil for postoperative analgesia. A pharmacokinetic and pharmacodynamic evaluation. *Anesthesiology*, 79, 481–492.
- Werner, M. U., Duun, P., & Kehlet, H. (2004). Prediction of postoperative pain by preoperative nociceptive responses to heat stimulation. *Anesthesiology*, 100, 115–119.
- Yarnitsky, D., Simone, D. A., Dotson, R. M., Cline, M. A., & Ochoa, J. L. (1992). Single C nociceptor responses and psychophysical parameters of evoked pain: effect of rate of rise of heat stimuli in humans. *Journal of Physiology*, 450, 581–592.
- Zondervan, K. T., Cardon, L. R., Kennedy, S. H., Martin, N. G., & Treloar, S. A. (2005). Multivariate genetic analysis of chronic pelvic pain and associated phenotypes. *Behavior Genetics*, 35, 177–188.