ACCP
American College of Clinical Pharmacology
Advancing Clinical Care through Pharmacology

A Systematic Review and Pooled Analysis of Select Safety Parameters Among Normal Healthy Volunteers Taking Placebo in Phase 1 Clinical Trials

The Journal of Clinical Pharmacology
2017, 57(9) 1079–1087

DOI: 10.1002/jcph.913

Tina C. Young, MS, DrPH[1], Subasree Srinivasan, MD, MPH[1]*, Marion L. Vetter, MD, MS[1]*, Venkat Sethuraman, PhD, MBA[1]*, Zubin Bhagwagar, MD, DPhil[2]*, Ricardo Zwirtes, MD[1], Premkumar Narasimhan, MS[1], Tilda Chuang, MS[1]*, and Brendan J. Smyth, MD, PhD[1]*

Abstract

A systematic review of the Bristol-Myers Squibb normal healthy volunteers (NHVs) database identified phase 1 trials that included NHVs administered placebo with the aim of characterizing normal inter- and intraindividual safety parameter variability. Twenty-five single and multiple ascending dose studies, median duration 28 (2 to 63) days, were included in the pooled analysis (355 NHVs). Laboratory evaluations, vital signs, electrocardiograms, and adverse events were assessed. The most commonly occurring adverse event was headache (28 [7.9%] NHVs; 519.5 events/100 person-years). During the dosing period (on placebo), evaluations showed 5.1 events/100 measures of alanine aminotransferase and 7.3 events/100 measures of creatine kinase 1× above the upper limit of normal. Alanine aminotransferase and creatine kinase elevations occurred in 28 (7.9%) and 39 (11.0%) NHVs, respectively; 105 (30.3%) NHVs had low and 46 (13.3%) had high diastolic blood pressure. This analysis may inform future study designs and provide a context for interpretation of safety signals in early phase clinical trials.

Keywords

pharmaceutical R&D, pharmacology, clinical trials, clinical research, clinical pharmacology

First-in-human trials—often single ascending dose (SAD) or multiple ascending dose (MAD) studies—typically include normal healthy volunteers (NHVs) administered either a placebo or an investigational drug. Inclusion criteria for NHVs in clinical trials include general, therapeutic area, and product-specific safety parameters. These typically include, but are not limited to, protocol-defined values for vital signs, physical examination, electrocardiograms (ECGs), and laboratory tests.

Select safety findings (eg, abnormal laboratory values or ECG findings) from early phase clinical studies in NHVs may result in the delay or termination of those studies and/or future clinical development programs. However, deviation in safety parameters from reference ranges or intervals may not necessarily be pathologic or clinically significant. Beyond intrinsic analytical variation and technical errors,[1,2] intraindividual (eg, activity level, food intake, and postural position) and interindividual (eg, sex, age, and body mass index) variation can influence laboratory tests.[3,4] In addition, recurrent resampling may lead to an increased likelihood of variability. Given the sources of variability, the probability of the occurrence of a given parameter with a value exceeding a predefined threshold will depend on the parameter itself and on the frequency and interval of observations.

Currently, there are few published studies of NHVs receiving placebo that provide incidence rates and variability of safety parameters in controlled settings.[5–15] Thus, available data that can aid decision making are limited. Pooling data from multiple phase 1 studies provides a larger population, compared with individual studies, in which to analyze safety parameters within a population of interest. The present systematic

[1] Bristol-Myers Squibb, Princeton, NJ, USA
[2] Yale University and Alexion Pharmaceuticals, New Haven, CT, USA

Submitted for publication 23 November 2016; accepted 14 March 2017.

Corresponding Author:
Tina C. Young, MS, DrPH, Bristol-Myers Squibb, Rt. 206 & Province Line Road, Princeton, NJ 08540-4000
E-mail: tina.young@bms.com

*Drs Zubin Bhagwagar, Venkat Sethuraman, Brendan Smyth, Subasree Srinivasan, and Marion L. Vetter and Tilda Chuang are former employees of Bristol-Myers Squibb.

review was performed on the Bristol-Myers Squibb (BMS) NHV database to facilitate safety review and decision making. The review identified phase 1 SAD and MAD studies that included NHVs taking placebo to determine the probability of select safety events based on predetermined criteria, encounter-adjusted incidence rates of select laboratory and ECG events, and exposure-adjusted adverse events (AEs).

Methods

Study Design and Analysis Population

This study included a systematic review to identify SAD and MAD studies within the BMS Exploratory Clinical and Translational Research database, conducted from January 1, 2010 through July 14, 2014, that included NHVs administered placebo. Drug–drug interaction and other proof-of-concept clinical pharmacological studies were excluded from this analysis, as subjects often receive other treatments along with placebo. The homogeneity of inclusion and exclusion criteria for each study was reviewed to determine whether they were appropriate for inclusion in a pooled analysis to evaluate the occurrence of select safety events based on predetermined criteria, both pretreatment (baseline) and during the dosing period (on placebo).

The analysis population included all subjects (NHVs) who signed an informed consent form, were enrolled into a phase 1 study, and who received at least 1 dose of placebo.

Study Selection Criteria

A catalog of all completed Exploratory Clinical and Translational Research studies, from January 1, 2010 through July 14, 2014, was compiled and evaluated for their appropriateness for inclusion in the analysis based on the following criteria: (1) the protocol had to have been approved and have a protocol number; (2) the population must consist only of NHVs, as defined within each study's protocol inclusion/exclusion criteria; (3) eligible studies must have an actual placebo arm; (4) the study type must have been of an SAD, MAD, or a combination of SAD/MAD design; and (5) the trial must have had a locked database. Studies containing patients or subjects of unknown patient status were excluded, as were studies that did not have a placebo arm or that used a cointerventional drug in combination with placebo and oncology studies.

Inclusion and exclusion criteria were applied by 3 independent reviewers (T.C.Y., S.T., and Z.B.) to ensure interrater reliability. The final selection of studies was reviewed (by T.C.Y., S.T., B.J.S., M.L.V., and S.S.) for similarity of design and collection schedule. The following attributes were assessed: (1) number of placebo subjects; (2) trial duration; (3) similar inclusion/exclusion criteria for healthy subjects (including age and smoking status); (4) similar data structures; and (5) common laboratory units (International System of Units, US, or both).

Safety Evaluation Criteria

The assessment of safety parameters was based on the following measurements for NHVs at baseline and during the dosing period (on placebo) from SAD, MAD, or SAD/MAD studies included in this analysis: specified laboratory evaluations of interest including liver tests (alkaline phosphatase, alanine aminotransferase [ALT], aspartate aminotransferase [AST], and total bilirubin), renal function tests (blood urea nitrogen and creatinine), pancreatic enzymes (amylase and lipase), creatine kinase (CK), and blood cell counts (including hemoglobin); vital signs including heart rate, systolic blood pressure (SBP), and diastolic blood pressure (DBP); and ECG. Subjects meeting the safety signal criteria in these measurements were also recorded.

Potential safety signals for the majority of laboratory evaluations were identified based on values that fell outside of normal reference ranges, as defined by the laboratory conducting the assays for each individual study. For hematology laboratory evaluations, potential safety signals were based on ranges defined in the BMS Global Standard documentation, Marked Abnormality Criteria For Clinical Laboratory Tests in Exploratory Clinical and Translational Research Studies in healthy subjects (Supplementary Table S1). In addition, hepatic function was specifically assessed using the following criteria defined in the US Food and Drug Administration Guidance for Industry on Drug-Induced Liver Injury:[16] ALT or AST $>3 \times$ upper limit of normal (ULN), $>5 \times$ ULN, >10 times ULN, and $>20 \times$ ULN; total bilirubin >2 times ULN; and concurrent ALT or AST $>3 \times$ ULN and total bilirubin $>2 \times$ ULN.

Safety signal criteria for vital signs were defined as follows: heart rate <40 beats per minute (low) and >100 beats per minute (high); SBP <90 mm Hg (low) and >160 mm Hg (high); and DBP <60 mm Hg (low) and >90 mm Hg (high). Safety signal criteria for ECGs were based on the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use E14 Guidance and defined as follows for QT prolongation:[17] >450 milliseconds for men, >470 milliseconds for women; >500 milliseconds "marked QT/QTc prolongation"; and change from baseline of >30 milliseconds and >60 milliseconds.

AEs occurring during the course of each study were recorded and included in this analysis. Serious AEs (SAEs), including deaths, were determined by the investigator(s) for each study.

Statistical Analysis

All analyses were performed using the statistical software SAS (version 9.2 or higher; SAS Institute Inc, Cary, North Carolina). Because this study was not designed to test a specific hypothesis, the sample size was not determined based on statistical power calculations. It was estimated (prior to the selection of data) based on the precision of the estimate and yielded a range of 63 to 323 NHVs in order to see specific effects (eg, ECG or liver tests) based on event rates seen in placebo subjects from published literature. For ALT > 1 time > ULN, Patat et al[18] provided several estimates of prevalence across multiple studies, and a conservative estimate of 22% was selected; a 95%CI covering an expected proportion of 0.22 with 5% precision would require 264 subjects. For QT prolongation, the estimate for >450 milliseconds, 6/649 (0.92%) indicates that it is a rare event, given that this is a population of NHVs; thus, a 95%CI covering the observation of ≥1 event with a true population rate of 0.00923 would require 323 subjects.[14] Sibille et al[19] considered 24 phase 1 studies in 430 healthy volunteers and found 20 "severe" AEs; to observe a 95%CI with this rate would require 63 subjects. The number of NHVs included in the analysis was 355, although certain analyses were limited due to availability of data and, therefore, have fewer subjects.

Baseline values were defined as the last value prior to the first dose of study drug (placebo). Categorical variables and binary event rates were summarized by frequencies and percentages. Event incidence rates, adjusting for subject encounters, among pooled placebo subjects were calculated for laboratory parameters, vital signs, and ECGs per 100 subject-encounters. Subject encounters were calculated for each safety parameter separately. Prevalence rates and corresponding CIs for safety parameters were calculated using a generalized linear mixed-effects model to take into account repeated measures within a subject. Continuous lab and ECG measurements were modeled using a linear repeated-measures model. Covariance structures used were compound symmetry and/or autoregressive(1) covariance structure (appropriateness of covariance structures were first assessed and differ by parameter). AEs were measured per person-year. Incidence rates were assumed to follow a Poisson distribution, and CIs were calculated using the method described by Haenszel et al.[20] For subject-level analyses, only the highest or lowest value for each subject was used depending on the criteria (eg, the highest value for drug-induced liver injury, the lowest value for neutropenia). The Rule of Three formula[21] was used to determine the population-level incidence rate if no events meeting the safety signal criteria were observed in the study sample.

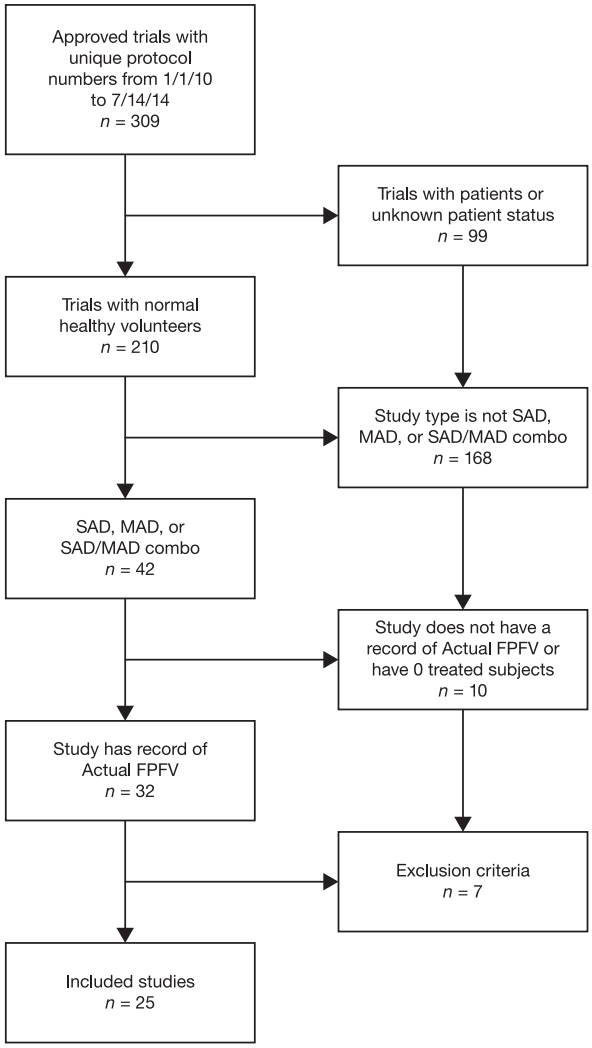

Figure 1. Study selection process for inclusion in the pooled analysis. The number of trials yielded by the search strategy used in the systematic review. Studies and data included in the analysis were sourced from the Management Analysis and Reporting System Exploratory Clinical and Translational Research studies from 2010 to 2014, generated July 14, 2014. FPFV indicates first patient first visit; MAD, multiple ascending dose; SAD, single ascending dose.

Results

Study Selection

A total of 25 phase 1 SAD and/or MAD trials were identified for inclusion in the pooled analysis. A flow diagram of the selection process is shown in Figure 1. These 25 trials included a total of 355 NHVs who had taken placebo during a study from multiple therapeutic areas, including virology, neuroscience, immunoscience, cardiovascular, and metabolic disease. The current analysis included studies carried out in 5 countries, including at least 8 different sites in the United States, which allowed for a diverse NHV population across geographic regions and countries. The duration of studies ranged from 2 to 63 (median 28) days; there

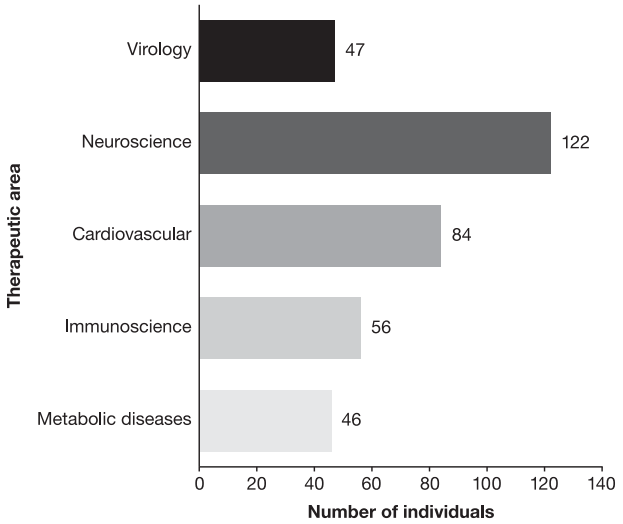

Figure 2. Therapy areas included in the pooled analysis. A total of 25 studies were included in the analysis: 4 virology, 9 neuroscience, 6 cardiovascular, 3 immunoscience, 3 metabolic disease, and no oncology studies. Three studies enrolled 0 to 5 placebo subjects, 7 studies enrolled 6 to 10 placebo subjects, 6 studies enrolled 11 to 15 placebo subjects, 4 studies enrolled 16 to 20 placebo subjects, and 5 studies enrolled more than 20 placebo subjects.

were 4 trials with durations of ≤10 days and 21 trials of longer duration. Different contract research organizations were used to conduct the studies included in the analysis.

The therapeutic areas with the largest number of studies and subjects were neuroscience and cardiovascular disease; approximately one-third of the subjects included in the analysis had participated in neuroscience studies (Figure 2). Because phase 1 oncology studies are typically performed neither with a placebo control nor in NHVs, no oncology studies were included. The distribution of therapeutic areas of the studies in the present analyses reflects the company's research program at the time of conduct.

Subject Demographic Data and Baseline Characteristics

Trial inclusion and exclusion criteria are summarized in Supplementary Table S2. The demographic data and baseline characteristics of NHVs included in the selected studies are shown in Table 1.

The study population consisted mainly of white (61%) American (73%) males (80%), of whom 59% were between 25 and 44 years old. However, 5% of the population was older than 65 years, as some NHVs included in the neuroscience therapeutic area were elderly (65 to 79 years), although otherwise healthy, and all met the criteria for inclusion at their trial site.

The median body mass index was 26 kg/m^2 (interquartile range 23 to 29 kg/m^2), indicating a slightly overweight study population, likely due to the fact that many of the subjects included in studies from the metabolic therapeutic area were obese (body mass index >30 kg/m^2) although otherwise healthy.

Table 1. Patient Demographic Data and Baseline Characteristics for Studies Included in the Pooled Analysis

	Placebo (N = 355)
Age, years	
Mean ± SD	36.0 ± 13
Median (range)	34 (18, 79)
Age category, years, n (%)	
<25	68 (19)
25-34	119 (34)
35-44	92 (26)
45-65	57 (16)
>65	19 (5)
Sex, n (%)	
Male	284 (80)
Female	71 (20)
Race, n (%)	
White	216 (61)
Black/African American	104 (29)
Asian	13 (4)
Native Hawaiian/Other Pacific Islander	1 (<1)
Chinese	13 (4)
Japanese	4 (1)
Other	4 (1)
Ethnicity, n (%)	
Hispanic/Latino	73 (21)
Not Hispanic/Latino	148 (42)
Not reported	134 (38)
Geographic distribution, n (%)	
Australia	52 (15)
China	5 (1)
Germany	28 (8)
Sweden	11 (3)
United States	259 (73)
Baseline BMI (kg/m^2)	
Mean ± SD	26 ± 4
Median (range)	26 (17, 40)
Baseline heart rate (bpm)	
Mean ± SD	65 ± 11
Median (range)	65 (42, 97)
Baseline systolic blood pressure (mm Hg)	
Mean ± SD	118 ± 12
Median (range)	117 (93, 167)
Baseline diastolic blood pressure (mm Hg)	
Mean ± SD	74 ± 9
Median (range)	73 (50, 95)
Baseline QTcF prolongation (milliseconds)	
Mean ± SD	402 ± 17
Median (range)	404 (351, 443)

BMI indicates body mass index; bpm, beats per minute; QTcF, QT interval corrected for heart rate using the Fridericia formula.

For most subjects (99.6%) included in the analysis, baseline heart rate, SBP, and QT interval (corrected for heart rate using Fridericia's formula [QTcF]) appeared normal. These findings were as expected for a population of NHVs accepted into each trial.

Adverse Events

AEs occurring in ≥2% of subjects are shown in Table 2. The most commonly occurring class of

Table 2. Summary of Adverse Events Experienced by ≥2% of Subjects During the Dosing Period

Safety Parameter	Number (%) of Patients With Event (N = 355)	Total Number of Events (Total PY = 6.93)	IR/100 PY	95%CI[a]
Headache	28 (7.9)	36	519.5	362.1, 722.1
Constipation	15 (4.2)	17	245.3	143.0, 392.5
Abdominal pain	7 (2.0)	7	101.0	40.5, 208.1
Upper respiratory tract infection	7 (2.0)	7	101.0	40.5, 208.1

IR indicates incidence rate; PY, person-years.
[a]CIs for incidence rates >1/100 were calculated using the method described by Haenszel et al.[21]

AEs were nervous system disorders and gastrointestinal disorders. The most commonly occurring AE was headache (28 [7.9%] subjects; 519.5 events/100 person-years), followed by constipation, abdominal pain, and upper respiratory tract infections. No SAEs, deaths, or events qualifying for the drug-induced liver injury criteria were observed, as would be expected for this population. Using the Rule of Three formula,[21] there was 95% confidence that fewer than 3 in 355 subjects taking placebo as a healthy subject would experience an SAE or death in a phase 1 SAD/MAD trial.

Laboratory Evaluations

As expected for NHVs, most laboratory results for liver function, renal function, and blood cell counts were within normal ranges at baseline and during the dosing period (on placebo). ALT and CK elevations were the most common deviations from safety parameter reference ranges (Table 3; Supplementary Table S3). During the dosing period, safety evaluations showed 5.1 events/100 measures of ALT >ULN (95%CI 3.0, 6.6) and 7.3 events/100 measures of CK >ULN (95%CI 6.1, 10.5), accounting for repeated measures among subjects. ALT levels increased to >1 × ULN to ≤2 × ULN and to >2 × ULN in 26 (7.3%) and 2 (0.6%) subjects, respectively, who had ALT levels below ULN at baseline. CK levels increased to >1 × ULN to ≤2 × ULN and to >2 × ULN in 31 (8.7%) and 8 (2.3%) subjects, respectively, who had CK levels below ULN at baseline. Elevations >ULN for alkaline phosphatase, AST, total bilirubin, blood urea nitrogen, creatinine, hemoglobin, and blood cell counts were observed only in very small numbers of subjects (Table 3; Figure 3; Supplementary Table S3).

During the dosing period, scatter plots for values for peak (highest) total bilirubin compared with ALT (Supplementary Figure S1) and AST (data not shown) for all placebo subjects showed no value for any subject that met safety signal (potential drug-induced liver injury) criteria. When hematology laboratory evaluations were compared with safety signal criteria for laboratory tests in Exploratory Clinical and Translational Research studies in healthy subjects (Supplementary Table S1), only 2 subjects demonstrated baseline values that met those criteria: 1 had a low lymphocyte cell count, and the other 1 had a high eosinophil cell count (Supplementary Table S4). Neutropenia was diagnosed in 15 subjects (4.2%) postbaseline (Supplementary Table S4), of whom 11 were black/African American and 4 were white. As described in detail in the Discussion, ethnic variations in neutropenia are common.

Vital Signs

As expected for NHVs, relatively few vital sign recordings met safety signal criteria at baseline (Supplementary Table S4). At baseline, no subject had a heart rate meeting the safety signal criteria, only 1 (0.4%) subject had high SBP, while 17 (6.2%) subjects had low DBP, and 7 (2.6%) subjects had high DBP. While receiving placebo, 8 (2.3%) subjects had a high heart rate; 14 (4.0%) subjects had low SBP, and 3 (0.8%) had high SBP; 105 (30.3%) subjects had low DBP, and 46 (13.3%) had high DBP.

Electrocardiogram

As expected in this NHV population, no subject experienced QTcF prolongation at baseline. Correlation was found to be the same among all pairs of observations within a subject. Females had an average QTcF 10.4 milliseconds higher than males ($P < .01$). Mean change from baseline over time was –0.22 milliseconds (95%CI –1.16, 0.72), taking into account repeated measures within a subject. There were 27 subjects experiencing a change in QTcF >30 milliseconds and 2 by >60 milliseconds. The rate of change in QTcF >30 milliseconds is 1.1 events/100 measures (95%CI 0.8, 1.5), accounting for repeated measures among subjects. While on placebo, 3 male subjects experienced a QTcF >450 milliseconds (Table 3), ranging from 451 to 473 milliseconds. One male had 1 elevation of 472 milliseconds, another 1 had 4 elevations within the same day (464, 465, 451, and 465 milliseconds), and the third male subject had 3 elevations within the same day (473, 457, and 453 milliseconds) and then again 20 days later (460 and 461 milliseconds). One female subject experienced a QTcF >470 milliseconds (483 milliseconds) (Supplementary Table S4). QTcF prolongation >500

Table 3. Summary of Safety Evaluations and Events Meeting the Safety Signal Criteria During the Dosing Period

Safety Parameter	N[a]	Total Number of Measures[b]	Number/Event	Total Number of Events	Number of Events/100 Measures	Model-Based IR and 95%CI[c]
ALT >1× ULN	355	1,587	32	81	5.1	4.5 (3.0, 6.6)
AST >1× ULN	355	1,574	9	13	0.8	NC
CK >1× ULN	323	1,478	61	108	7.3	8.0 (6.1, 10.5)
TBil >1× ULN	355	1,563	9	10	0.6	NC
QTcF (male) >450 milliseconds	267	3,423	3	10	0.3	NC
QTcF (female) >470 milliseconds	69	1,412	1	1	0.1	NC
Change in QTcF >30 milliseconds	336	4,835	27	53	1.1	1.1 (0.8, 1.5)
Change in QTcF >60 milliseconds	336	4,835	2	4	0.1	NC
Creatinine >ULN	337	1,496	7	16	1.1	NC

ALT indicates alanine aminotransferase; AST, aspartate aminotransferase; CK, creatine kinase; IR, incidence rate; NC, not calculated; QTcF, QT interval corrected for heart rate using the Fridericia formula; TBil, total bilirubin; ULN, upper limit of normal.

[a]Number of placebo subjects with postbaseline measurements.

[b]Postbaseline measurements.

[c]CIs for parameters with N > 10 were calculated from the generalized linear mixed-effects model.

milliseconds was not observed in any subject either at baseline or during the dosing period. A change in QTcF from baseline >30 milliseconds was reported in 27 subjects (7.6%), and 2 subjects (0.6%) experienced a change from baseline of >60 milliseconds (Supplementary Figure S2; Table 3). Of the 27 subjects with a change in QTcF from baseline >30 milliseconds, 24 were 18 to 50 years old, and the other 3 were >70 years old. Of those 27 subjects, the 2 with a change in QTcF >60 milliseconds were a 75-year-old female and a 19-year-old male.

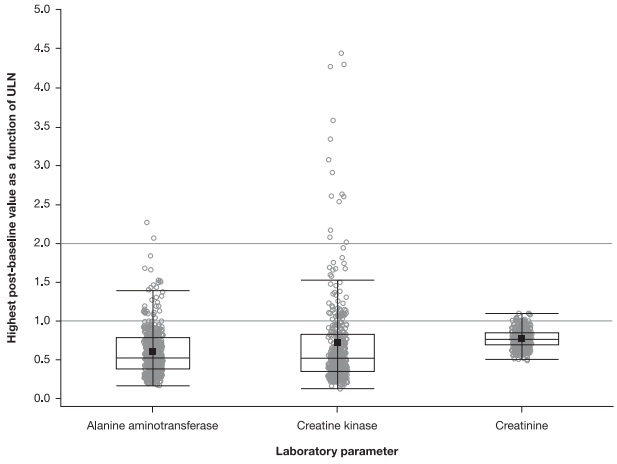

Figure 3. Highest postbaseline values of alanine aminotransferase, creatine kinase, and creatine as a function of ULN. One NHV had creatine kinase 30 × ULN and was excluded from this plot. The box plot shows median (middle line) and interquartile range. The whiskers show 1.5 × interquartile range. NHV indicates normal healthy volunteer; ULN, upper limit of normal.

Discussion

This systematic review and pooled analysis of 25 SAD/MAD randomized controlled studies (median duration 28 days) from within the BMS study database included 355 NHVs and identified relatively few changes in laboratory parameters or vital signs that met the study safety signal criteria. During the dosing period (on placebo) evaluations showed 5.1 events/100 measures of ALT and 7.3 events/100 measures of CK >ULN. There was some variation in DBP: 105 (30.3%) NHVs had low DBP and 46 (13.3%) had high DBP across all studies. These pooled placebo data could help inform study investigators when interpreting potential safety signals observed in phase 1 and later clinical trials by providing a reference point or comparison for subjects on active treatment.

As expected in a population of NHVs, the majority of laboratory results were within normal reference ranges at baseline and during placebo treatment.[16] ALT elevation occurred in 28 (7.9%) NHVs. Previous retrospective pooled analyses of transaminase elevation in the placebo arms of 13 phase 1 clinical trials with male NHVs (n = 93) found that 20.4% had at least 1 ALT value exceeding the ULN and 7.5% had at least 1 ALT value exceeding twice the ULN.[8] The authors noted that transaminase elevation in NHVs during phase 1 clinical trials is a real phenomenon, explained largely if not completely by multiplicity of sampling, prompting the cautious interpretation of laboratory safety results, in particular hepatotoxicity.[8] In 974 adult NHVs who participated in clinical trials at the Jenner Institute, Oxford, UK, between 1999 and 2009, among a range of tested analytes, ALT, total bilirubin, white blood cells, and urea showed the greatest intraperson variation.[10] Of note, age, body mass index, and sex are associated with changes in ALT.[10,22,23]

CK elevation occurred in 39 (11.0%) NHVs in our study. CK elevation has been observed in response to exercise or muscle trauma.[24,25] Natural and/or ethnic variations in serum CK levels may occur in the general

population, which can further confound interpretation. Evaluation of reference intervals for serum CK levels in a stratified random sample of 1444 individuals from the general population in the Netherlands, according to National Committee on Clinical Laboratory Standards/Nordic Reference Interval Project guidelines, showed that 13% of white Europeans, 23% of South Asians, and 49% of black people in the sample had serum CK levels above manufacturer-provided limits.[26] Furthermore, there is evidence to show that healthy black people have CK levels around 70% higher than white people.[27]

Hepatotoxicity is a leading reason for halting clinical development of new chemical entities. Thus, careful interpretation of safety findings, including data from subjects taking placebo, is critical to decision making and future safety monitoring in clinical trials. Importantly, the more frequently a measure is sampled, the higher is the likelihood of observing a value that deviates from the normal range. This may particularly be the case in phase 1 studies, where sampling of laboratory measures, vital signs, and ECG is very frequent, which is important to take into consideration when evaluating values that meet safety signal criteria. Additionally, the inclusion and exclusion criteria for phase 1 clinical trial populations need to reflect the therapeutic area under study, as effects in the NHV population will be directly relevant to subsequent phase 2/3 trial populations.

In the present analysis, 15 (4.2%) subjects had neutropenia postbaseline during the course of the study in which they were participating, 11 of whom were black/African Americans. Benign ethnic neutropenia is neutropenia (generally neutrophil counts $<1.5 \times 10^9$ cells/L) in otherwise healthy individuals and has been observed at higher rates in different ethnic groups including Africans, Afro-Caribbeans, and black/African Americans.[28,29] It is not known whether the cases of neutropenia in the present study were related to benign ethnic neutropenia. Nevertheless, ethnic variations in neutropenia are common, and it has been suggested that rigid requirement for minimal white blood cell numbers may result in individuals with benign ethnic neutropenia being excluded from participating in clinical trials.[28]

Low DBP was recorded in 30.3% of NHVs on placebo in the present study. Multiple factors are known to affect blood pressure readings.[30] Consistent with statistics from the Centers for Disease Control and Prevention, high DBP was more likely in older NHVs, and low DBP was more likely in younger NHVs.[31] As previously reported, sex was also associated with trends in DBP: high DBP was more common in men and low DBP was more common in women.[32] The American Heart Association reported that high DBP was more common in men than in women in the <45-year age group, but this trend was reversed in the >65-year age group.[33] These reports are consistent with the NHVs included in our analysis, approximately 80% of whom were aged <45 years. High DBP was more common in metabolic disease trials compared with other therapeutic areas, possibly due to the eligibility criteria permitting NHVs with higher body mass index,[34] and NHVs enrolled in cardiovascular disease trials were less likely to have low DBP. Interestingly, if blood pressure criteria were updated based on values derived from the normal distribution, using the mean ± 1.96 times the standard deviation, the definition of low DBP would be revised from <60 mm Hg to <52 mm Hg, decreasing the incidence of low DBP from 30.3% to 7.8%. These findings are highly relevant to phase 1 trials of therapeutic agents with cardiovascular effects, and the inclusion and exclusion criteria for such trials should therefore be carefully considered.

The absence of QTcF prolongation at baseline reflects the study exclusion criterion QTcF >450 milliseconds for 24 of the 25 phase 1 SAD/MAD trials included. The majority of ECG results meeting the safety signal criteria occurred in cardiovascular studies, which may have arisen due to the increased frequency of ECG sampling that took place in studies conducted in neuroscience and cardiovascular disease compared with other therapeutic areas.

Data for 844 NHVs from 9 thorough QT studies were used to compare the effects of active control (moxifloxacin) and placebo on QTcF. QT intervals meeting the safety signal criteria were rare in both the placebo and active treatment groups: QTcF >450 milliseconds occurred in 2.9% and 0.9% of moxifloxacin and placebo groups, respectively, with no cases of QTcF >480 milliseconds in either group. Intra- and interindividual variations (standard deviation) in QTcF were approximately 8 to 9 milliseconds and 9 to 10 milliseconds, respectively.[14] The QTcF prolongation results of the current analysis for subjects taking placebo were similar to these previous reports. The mean change from baseline in QTcF over time was –0.22 milliseconds (95%CI –1.16, 0.72), which is in agreement with the range of mean change from baseline (–16.3 to 2.9) reported in a retrospective analysis of placebo and nondrug data from ~380 patients included in 5 clinical trials.[35]

Pooled analyses assessing safety parameters in phase 1 clinical trials in NHVs have been reported previously; however, these mainly focus on the AE profile for placebo vs investigational drugs. In a survey of adult NHVs who participated in phase 1 clinical trials at the Bayer Health Care AG in Wuppertal, Germany, between 1994 and 1999 (1559 NHVs from 142 studies; included 3862 follow-up days on placebo), there were 351 AEs in NHVs assigned to placebo (301 of mild and 49 of moderate severity). The most commonly

occurring types of AEs were general disorders (3.0%) and headache (2.5%).[7] This analysis was extended to look at phase 1 dose-escalation studies in male NHVs (24 studies; 18 to 46 years).[15] A total of 368 NHVs received placebo and 82 (22.3%) experienced AEs (74 of mild and 8 of moderate severity). The most common AEs by system organ class were nervous system disorders (31 events), infections and infestations (18 events), and gastrointestinal disorders (15 events). Similar results were found in the present pooled analysis, where the most commonly occurring AE was headache (28 individuals). Importantly, there were no SAEs or deaths in any of the 25 trials included.

The sample size of the present study, 355 subjects in 25 studies, was relatively small when compared with similar published studies by other biopharmaceutical companies, including a recent report of 1559 NHVs from 142 studies and an analysis of 11,028 placebo subjects in 394 distinct nononcology phase 1 studies.[6,7] However, both of these previous studies used an all-comer selection criteria approach that included all nononcology phase 1 studies. In contrast, the present analysis employed very specific inclusion/exclusion criteria to answer the study objectives. Thus, only SAD, MAD, or combined SAD/MAD studies were included. Drug–drug interaction and proof-of-concept studies, for example, were excluded on the basis that they do not contain a true placebo arm. In addition, based on the sample size calculation, only 325 subjects were needed to answer our study objectives. The sample size was estimated based on the precision of the estimate in order to see specific effects based on the published literature.[14,18,19]

Limitations should be considered. This analysis included trials identified from the BMS phase 1 database. Study selection was limited to those studies and respective therapeutic areas within the database. All studies were funded by the same study sponsor (BMS); thus, homogeneity in data collection and trial practices can be expected, although these practices may differ from those of other clinical trial sponsors. Included trials spanned a range of therapeutic areas (Figure 2) and, as a result, different subject and site selection criteria may have applied. This analysis was not planned "per protocol" and should be considered as "observational" and therefore subject to limitations that apply to this type of study design. Of note, the AE profile under placebo may be influenced by active treatment as NHVs are informed about substance-specific AEs.[7] The median (range) study duration was 28 (2 to 63) days, which, although longer than reported in previous analyses of NHV studies (14 to 17 days),[7,15] is a relatively short period of observation.

The accessibility of accurate and appropriate safety parameters is critical for the proper screening of NHVs, determination of clinical trial inclusion and exclusion criteria appropriate to a given therapeutic area, and monitoring of NHVs during phase 1 studies and to inform dosing decisions. The availability of the BMS NHV phase 1 study database provides the opportunity to assign normal reference ranges for safety parameters based on placebo-controlled criteria that are relevant for drug development programs in each therapeutic area. Other potential useful applications of these data include informing designs for future clinical trials (such as Bayesian priors or the Δ for sample size calculations), the reduction in randomization to phase 1 trial placebo arms by using these data as historical controls, and use in epidemiology studies as a comparator arm.

Conclusion

In summary, there were no safety concerns involving laboratory parameters, ECGs, SAEs, or deaths. DBP measurements meeting the safety signal criteria were as noted. Safety parameters in phase 1 NHV studies can deviate from the normal range, possibly due to multiplicity of sampling and fluctuations in normal physiology. The safety parameter out-of-range findings in NHVs receiving placebo presented here may serve as a reference to strengthen the basis for signal detection and decision making during early drug development, and inclusion and exclusion criteria for future studies in NHVs should be reexamined with these results in mind.

Acknowledgments

Professional medical writing and editorial assistance was provided by Catriona McKay and Stephen Moore at Caudex and was funded by Bristol-Myers Squibb. Zemin Zeng (BMS) contributed to the construction of the pooled database and provided programming support for the analysis. Susan Takacs contributed to the review and selection of studies for inclusion in this systematic review and pooled analysis, and Naveen Talluri provided programming support for the analysis; this work was funded by Bristol-Myers Squibb.

Declaration of Conflicting Interests

T.C.Y., R.Z., and P.N. are employees of Bristol-Myers Squibb. Z.B., V.S., B.J.S., S.S., M.L.V., and T.C. are former employees of Bristol-Myers Squibb.

Funding

This study was sponsored by Bristol-Myers Squibb.

Author Contributions

All authors contributed to the writing of the manuscript, designed and carried out the study research, and analyzed the data.

References

1. Carraro P, Plebani M. Errors in a stat laboratory: types and frequencies 10 years later. *Clin Chem*. 2007;53:1338–1342.
2. Plebani M, Carraro P. Mistakes in a stat laboratory: types and frequency. *Clin Chem*. 1997;43:1348–1351.
3. Phillips P. Pitfalls in interpreting laboratory results. *Australian Prescriber*. 2009;32:43–46.
4. Henny J, Petitclerc C, Fuentes-Arderiu X, et al. Need for revisiting the concept of reference values. *Clin Chem Lab Med*. 2000;38:589–595.
5. Ceriotti F, Henny J, Queralto J, et al. Common reference intervals for aspartate aminotransferase (AST), alanine aminotransferase (ALT) and gamma-glutamyl transferase (GGT) in serum: results from an IFCC multicenter study. *Clin Chem Lab Med*. 2010;48:1593–1601.
6. Emanuel EJ, Bedarida G, Macci K, Gabler NB, Rid A, Wendler D. Quantifying the risks of non-oncology phase I research in healthy volunteers: meta-analysis of phase I studies. *BMJ*. 2015;350:h3271.
7. Lutfullin A, Kuhlmann J, Wensing G. Adverse events in volunteers participating in phase I clinical trials: a single-center five-year survey in 1,559 subjects. *Int J Clin Pharmacol Ther*. 2005;43:217–226.
8. Rosenzweig P, Miget N, Brohier S. Transaminase elevation on placebo during phase I trials: prevalence and significance. *Br J Clin Pharmacol*. 1999;48:19–23.
9. Roux C, Chevassus H, Farret A, Costa F, Petit P, Galtier F. Selecting healthy volunteers in specific populations: a retrospective analysis of clinical and laboratory screening. *Fundam Clin Pharmacol*. 2010;24:121–127.
10. Rowland R, O'Hara GA, Hamill M, et al. Determining the validity of hospital laboratory reference intervals for healthy young adults participating in early clinical trials of candidate vaccines. *Hum Vaccin Immunother*. 2013;9:1741–1751.
11. Sibille M, Lassonery LG, Janin A, Deigat N, Boutouyrie B, Durand DV. Upper limit of plasma alanine amino transferase during phase I studies. *Eur J Clin Pharmacol*. 1995;47:417–421.
12. Sibille M, Bresson V, Janin A, Boutouyrie B, Rey J, Vital DD. Critical limits to define a lab adverse event during phase I studies: a study in 1134 subjects. *Eur J Clin Pharmacol*. 1997;52:81–86.
13. Sibille M, Deigat N, Durieu I, et al. Laboratory data in healthy volunteers: reference values, reference changes, screening and laboratory adverse event limits in Phase I clinical trials. *Eur J Clin Pharmacol*. 1999;55:13–19.
14. Stylianou A, Roger J, Stephens K. A statistical assessment of QT data following placebo and moxifloxacin dosing in thorough QT studies. *J Biopharm Stat*. 2008;18:502–516.
15. Wensing G, Ighrayeb IA, Boix O, Böttcher M. The safety of healthy volunteers in first-in-man trials—an analysis of studies conducted at the Bayer in-house ward from 2000 to 2005. *Int J Clin Pharmacol Ther*. 2010;48:563–570.
16. U.S. Department of Health and Human Services Food and Drug Administration. Guidance for industry: drug-induced liver injury: premarketing clinical evaluation. 2009. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM174090.pdf.
17. U.S. Department of Health and Human Services Food and Drug Administration. ICH E14—The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. 2005. http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073153.pdf.
18. Patat AA. Designing and interpreting the results of first-time-to-man studies. *Dialogues Clin Neurosci*. 2000;2:203–212.
19. Sibille M, Deigat N, Olagnier V, Durand DV, Levrat R. Adverse events in phase one studies: a study in 430 healthy volunteers. *Eur J Clin Pharmacol*. 1992;42:389–393.
20. Haenszel W, Loveland DB, Sirken MG. Lung-cancer mortality as related to residence and smoking histories. I. White males. *J Natl Cancer Inst*. 1962;28:947–1001.
21. Eypasch E, Lefering R, Kum CK, Troidl H. Probability of adverse events that have not yet occurred: a statistical reminder. *BMJ*. 1995;311:619–620.
22. Prati D, Taioli E, Zanella A, et al. Updated definitions of healthy ranges for serum alanine aminotransferase levels. *Ann Intern Med*. 2002;137:1–10.
23. Salvaggio A, Periti M, Miano L, Tavanelli M, Marzorati D. Body mass index and liver enzyme activity in serum. *Clin Chem*. 1991;37:720–723.
24. Deminice R, Trindade CS, Degiovanni GC, et al. Oxidative stress biomarkers response to high intensity interval training and relation to performance in competitive swimmers. *J Sports Med Phys Fitness*. 2010;50:356–362.
25. Lippi G, Schena F, Montagnana M, Salvagno GL, Banfi G, Guidi GC. Significant variation of traditional markers of liver injury after a half-marathon run. *Eur J Intern Med*. 2011;22:e36–e38.
26. Brewster LM, Mairuhu G, Sturk A, van Montfrans GA. Distribution of creatine kinase in the general population: implications for statin therapy. *Am Heart J*. 2007;154:655–661.
27. Brewster LM, Coronel CM, Sluiter W, Clark JF, van Montfrans GA. Ethnic differences in tissue creatine kinase activity: an observational study. *PLoS One*. 2012;7:e32471.
28. Hsieh MM, Tisdale JF, Rodgers GP, Young NS, Trimble EL, Little RF. Neutrophil count in African Americans: lowering the target cutoff to initiate or resume chemotherapy? *J Clin Oncol*. 2010;28:1633–1637.
29. Paz Z, Nails M, Ziv E. The genetics of benign neutropenia. *Isr Med Assoc J*. 2011;13:625–629.
30. Frese EM, Fick A, Sadowsky HS. Blood pressure measurement guidelines for physical therapists. *Cardiopulm Phys Ther J*. 2011;22:5–12.
31. National Center for Chronic Disease Prevention and Health Promotion DfHDaSP. High blood pressure. 2015. http://www.cdc.gov/bloodpressure/facts.htm. Accessed October 21, 2015.
32. Herbert A, Cruickshank JK, Laurent S, Boutouyrie P. Establishing reference values for central blood pressure and its amplification in a general healthy population and according to cardiovascular risk factors. *Eur Heart J*. 2014;35:3122-3133.
33. Mozaffarian D, Benjamin EJ, Go AS, et al. Heart disease and stroke statistics—2015 update: a report from the American Heart Association. *Circulation*. 2015;131:e29–322.
34. Frohlich ED. Clinical management of the obese hypertensive patient. *Cardiol Rev*. 2002;10:127–138.
35. Agin MA, Aronstein WS, Ferber G, Geraldes MC, Locke C, Sager P. QT/QTc prolongation in placebo-treated subjects: a PhRMA collaborative data analysis. *J Biopharm Stat*. 2008;18:408–426.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website.