Rogers's RS und SC Malingering Scales Derived from the SIMS

Zack Z. Cernovsky[1]*, Jack Remo Ferrari[2]

[1]Department of Psychiatry, Western University, London, Ontario, Canada.

[2]Psychological Clinic, London, Ontario, Canada.

zcernovs@uwo.ca

Corresponding Author: *Zack Cernovsky, Dept. of Psychiatry, Western University, London, Ontario, Canada.*

Abstract

Background: *Recent studies have shown conclusively that the Structured Inventory of Malingered Symptomatology (SIMS) lacks in content validity and criterion validity. Rogers, Robinson, and Gillard (2014) used an innovative statistical procedure to extract those SIMS items that could still usefully differentiate legitimate psychiatric patients instructed to respond honestly from those instructed to exaggerate symptoms. Their procedure resulted in their Rare Symptoms (RS) scale and Symptom Combination (SC) scale. The present study applied the RS and SC scales to patients injured in high impact motor vehicle accidents (MVAs) and also critically evaluates the content of these scales.*

Method: *An ANOVA was calculated to compare RS and SC data of 3 groups: (1) 23 survivors of high impact MVAs, (2) data collected by Rogers's team on their 54 psychiatric patients instructed to respond honestly, and (3) data collected by Rogers's team on their 53 psychiatric patients instructed to exaggerate symptoms.*

Results and Discussion: *All except four of the 23 post-MVA patients, i.e., 82.6%, obtained scores below Rogers's cutoff of > 6, i.e., at a "non-malingering" level, on the RS scale and all except one of the 23, i.e., 95.7% scored below Rogers's cutoff of > 6, i.e., at a "non-malingering" level, on the SC scale. Damaging evidence against the RS scale comes from the ANOVA: the RS scores of post-MVA patients did not differ significantly from psychiatric patients instructed to exaggerate: both groups scored significantly higher than psychiatric patients responding honestly. The SC scores of post-MVA patients did not differ significantly in the ANOVA from psychiatric patients responding honestly: both groups scored significantly lower than psychiatric patients instructed to exaggerate. Clinical content analysis of RS scale suggests irremediable flaws. A third of the RS items are logical or algebraic reasoning tasks on which patients with severe post-concussive symptoms and fatigue from insomnia (such as caused by persistent pain) could perform less well. Patients with extensive microvascular injuries and axonal shearing from their accident are more likely to score higher on the RS and be misclassified as "malingerers" than less injured persons. Another third of RS scale items lists delusional symptoms or those of thought disorder: psychotic patients are more likely to be branded as "malingerers" and deprived of pharmacotherapy. The SC scale is based on a precarious assumption that correlations among its symptoms remain the same across varied groups of genuine medical patients, regardless of the type and intensity of their own symptoms. Patients more severely disabled by their symptoms might be less consistent in their responses and thus more often misclassified as "malingerers" by the SC.*

Furthermore, Rogers's psychiatric sample on which the RS and SC scales were developed was diagnostically mixed, too heterogeneous, mainly diagnosed with PTSD (>77%) and/or mood disorders (>32%): this makes generalizations of RS and SC cutoffs to other diagnostic groups of psychiatric patients uncertain. Generalizations to yet other medical patients would need to be tested carefully, separately for each diagnostic group and on larger samples to satisfy APA requirements.

Conclusions: *Too many items of the RS represent common rather than "Rare Symptoms." For many reasons, non-SIMS items would be a better choice for future RS scales. With respect to developing future SC scales, more extensive research would be needed to demonstrate if, or when, specific pairs of symptoms are indeed always mutually exclusive or uncorrelated.*

Keywords: *malingering, SIMS, rare symptoms scale, symptom combination scale, insurance claims*

INTRODUCTION

Richard Rogers is celebrated worldwide as expert for detection of malingering: his book "Clinical Assessment of Malingering and Deception"[1] is at its 4th edition. The present article evaluates statistical data on the Rare Symptoms (RS) and Symptom Combination (SC) scales published in 2014 by Rogers, Robinson, and Gillard.[2] Unfortunately, these RS and SC scales were derived solely from items of the Structured Inventory of Malingered Symptomatology (SIMS)[3] notorious for its lack of content validity[4,5,6,7] and criterion validity.[4,5,8,9] The SIMS has used seriously flawed cutoffs for its 6 scales that lead to high proportions of false positives even in normal control samples.[10]

Van Impelen, Mercklebach, Jelic, and Merten[11] indicated already in 2014 that SIMS has "*substandard specificity*" and expressed doubts about diagnostic utility of the cutoff of >14 points for SIMS total score, recommended by Glenn Smith (author of the SIMS) in his SIMS manual.[3]

Richard Rogers's team[2] mentioned already in 2014, with respect to SIMS cutoff of > 14 points, that "*research (e.g., Clegg et al., 2009*[12]*) has found that non-feigning clients often exceed this cut score.*" Rogers and his co-authors suggested in their 2014 article that the cutoff for SIMS total score might need to be set as high as > 44 points to improve specificity, when dealing with certain diagnostic groups. SIMS specificity calculated by Rogers et al.[2] for cutoff > 14 points was only .28.

In their 2014 article,[2] Rogers's team has developed an innovative methodological approach to examine statistically if some SIMS items could be extracted to form new scales to more successfully differentiate malingers from honest responders than do the classical SIMS scores, in particular with respect to feigned symptoms of mental illness. The participants in Rogers's 2014 study were described as patients of "*an adult inpatient trauma unit within the Timberlawn Mental Health System in Dallas, Texas. Thus, all patients*

had extensive trauma histories. They voluntarily participated in an intensive treatment program involving medications plus group and individual interventions."[2] These patients were randomly assigned to a group instructed to exaggerate their symptoms or to a group instructed to report their symptoms honestly, as follows. In the "feigned condition" the participants were asked to simulate total disability. For those currently already receiving disability, the instructions varied slightly. They were told the following "*I know you are currently disabled and cannot work. However, for this study, please imagine that your claim for disability payments was rejected. In other words, the insurance company did not think you were disabled enough to need any payments. Now imagine that you are going through another disability evaluation.*"[2]

The feigning group was given both internal (i.e., outsmarting the test) and external ($10) incentives for successfully simulating disability. The group was also cautioned that tests could include trick questions that were designed to detect persons faking the test. Finally, the importance of the research was stressed in the light of health care fraud, on the one hand, and wrongful denial of needed services because of suspected feigning on the other.

The genuine group was instructed to be forthright about their symptoms and impairment. They were told the following "*Please be open and honest about your current symptoms and psychological problems. It is important that you tell us the way it really is – please don't hide your problems or exaggerate your symptoms.*"[2] They were given the same external incentive ($10) for accurately disclosing their symptoms and other clinical characteristics.[2]

All participants were genuine mental patients. Thus, Rogers's study was to find SIMS items that would differentiate mental patients responding honestly from those mental patients who experience genuine mental illness, but also exaggerate their symptoms. Rogers used two strategies. The first was to locate

SIMS items listing medical symptoms reported rarely by the honest group but frequently by the exaggerating group: "*The rare symptoms (RS) scale was created by identifying SIMS items endorsed by less than 10% of genuine responders but more than 25% of feigners.*" The SIMS RS scale developed by Rogers contains 15 SIMS items.[2]

The second strategy used by Rogers was based on identifying unlikely combination of symptoms, those frequent in feigners but infrequent in the honest group. The following procedure was used to develop the SIMS symptom combination (SC) scale "*The correlations of all SIMS pairs were first calculated. Pairs of items were selected on two criteria: (a) they were uncorrelated or negatively correlated for genuine responders; and (b) they are positively correlated for feigners and accounted for more than 10% of the variance (φ coefficient >.35).*" [2] The SC scale contains 13 pairs of SIMS items.

Rogers's approach has probably indeed been the most intelligent attempt ever to transform the SIMS questionnaire from a pseudopsychological test into a legitimate diagnostic instrument.

One of the most serious methodological problems with these RS and SC scales lies in the diagnostically mixed nature of the sample of both groups of psychiatric patients recruited by Rogers's team: "*On average, the inpatient had been hospitalized at the treatment facility for just over a week (M= 7.81 days, SD = 5.36) before participating in the study. Nearly half of these patients (48% or 44.4%) had experienced multiple traumas, both childhood and adult, with large numbers having been diagnosed with post-traumatic stress disorder (83% or 77.5%). In addition, mood disorders predominated, with major depressive disorder (44% or 41.1%) and bipolar I disorder (35% or 32.7%) being the most common.*"[2]

If the RS and SC scale are to be used by clinicians, the test construction standards stipulated by the American Psychological Association (APA)[13] require that they should be used only on samples comparable to the one on which they were developed. For this reason, the criterion group on which the scale is developed and validated must be more adequately homogeneous and better defined. Rogers's sample consists of a heterogeneous mixture of patients, most of whom were traumatized, diagnosed with PTSD and/or mood disorders, presumably especially with depressive symptoms. Given the heterogeneous nature of these patients in Rogers's study, it is not clear whether the scales would perform well on other psychiatric samples such as, for instance, patients with psychoses in the spectrum of schizophrenia. Consistently with these APA requirements, Rogers's team cautioned the reader that their "*current findings may not generalize to other clinical and forensic populations.*"[2]

Obviously, more research on Rogers's RS and SC scales is needed to examine whether more homogeneous groups of psychiatric or other medical patients indeed differ significantly from instructed malingerers. Similarly, it is unclear whether RS and SC scores of patients injured in high impact motor vehicle accidents (MVAs) systematically differ from some groups of patients who exaggerate their symptoms: this is examined statistically in the present study. A thorough content analysis of the RS and SC scales is also required to determine if their items are prone to misclassify certain patients as "malingerers."

METHOD

Participants. Our sample of patients injured in high impact MVAs has been already described in a previous study[5] as follows "*23 survivors of high impact motor vehicle accidents (MVAs) in which their vehicle was damaged so extensively that it was subsequently deemed not worthy of repair. Such accidents are too rare without involving injuries such as those of a neuropsychological nature, especially symptoms in the post-concussion whiplash spectrum. The sample of our patients consists of 8 males and 15 females, age 19 to 60 years (mean age=38.0, SD=12.8), with education from 10 to 18 years (mean=14.1, SD=1.9). Their average scores were 17.2 (SD=11.0) on the Post-MVA Neurological Symptoms scale[14], 6.3 (SD=1.3) on the average pain item of the Brief Pain Inventory[15], and 23.7 (SD=3.0) on Morin's Insomnia Severity Index[16]. Their scores on the Insomnia Severity Index were known for 22 of the 23 patients: they were in Morin's categories of moderate insomnia for 6 patients (27.3%) and severe insomnia for 16 patients (72.7%). Such levels of insomnia are consistent with these patients' pain scores on the Brief Pain Inventory[15] because pain tends to disrupt sleep extensively. All patients in this sample could be classified as experiencing some degree of the post-concussion syndrome (scores ranging from 24 to 58 on the Rivermead scale[17], with mean=37.4, SD=13.2). The time elapsed since the patient's MVA ranged from 7 to 217 weeks, with the average at 81.5*

weeks (SD=55.8). However, all still experienced active post-accident symptoms. All still retained a lawyer to represent them to their car insurance company in disputes about payments for treatments and other benefits. The physical nature of their vehicular collision (high impact, with their car damaged to the extent of being deemed not worthy of repair) makes the accusation of malingering less plausible, even though some distressed patients may strongly emphasize their symptoms for fear of otherwise receiving no treatments or help." [5]

As already mentioned in the prior publication,[5] some insurance contracted psychologists may still erroneously assume that "*cerebral concussions occur too rarely without visible head injuries and without a complete and prolonged loss of consciousness. Neuropathological research by Bennet Omalu*[18, 19] *on players of American football demonstrated that cerebral damage in concussions occurs with sudden acceleration or deceleration of the head even in persons who neither sustained visible head injuries nor fully lost consciousness. These persons, within minutes after their concussion, may still be able to perform some simple physical tasks such as those involved in playing football. However, microvascular injuries and axonal shearing with subsequent neurotoxicity do occur in such incidents while the gray and the white parts of the brain slide over each other during the sudden excessive acceleration or deceleration of the skull.*" [5]

This study examines mean RS and SC scores of the 23 post-MVA patients and also compares their RS and SC scores to those of two groups of mental patients reported in the study led by Richard Rogers.[2] The data involving Rogers's RS and SC scales as well as all other SIMS scales were calculated in ANOVAs on these 3 groups:

(1) Rogers's group of 54 patients instructed to respond honestly.

(2) Rogers's group of 53 patients instructed to exaggerate their symptoms.

(3) Sample of 23 patients injured in high impact car accidents.

With respect to the 2 groups studied by Rogers's team, the mean age of these two groups (combined N=107) was 41.7 years (SD= 11.4). With respect to gender, 74.8% were females and 25.2% males. Their symptoms were mainly those of PTSD and of mood

disorder.[2] Unfortunately, these mental patients were not a diagnostically adequately homogeneous group of patients with exactly the same specific diagnosis.

Content of the RS and SC scales was analysed from a clinical perspective to determine whether certain diagnostic groups are at risk of being misclassified as "malingerers."

Results

Mean RS and SC Scores and Mean Other SIMS Scores

The mean RS and SC scores as well as all other SIMS scores of patients in Rogers's study[2] and of our 23 patients injured in high impact MVAs are listed in Table 1. The score cutoffs (those stipulated by SIMS manual and also those determined by Rogers's team for the RS and SC scales) are indicated in the bottom row of Table 1.

All these tabular values are rounded to one decimal point. It is noteworthy that, before rounding, the mean score on the SIMS LI scale of honest responders in Rogers's study was 2.02, i.e., above the cutoff > 2 points stipulated in SIMS manual for the LI scale. Thus, all 6 traditional SIMS mean scores (NI, AM, LI, AF, P, and the total score) of the genuine severely ill psychiatric patients in the group of 54 *honest responders* were above the cutoffs for malingering, see Table 1. The same is true about the mean scores on the 6 traditional SIMS scales of our 23 patients injured in high impact MVAs. Indeed, the SIMS, as developed by Glenn Smith, is a pseudopsychological test that misclassifies honest and legitimate patients as malingerers. If the normal distribution is assumed, mean scores above the cutoff usually imply that more than 50% of the samples might be misclassified by SIMS manual as "malingerers." The very fact that this occurred on all 6 traditional SIMS scales in both the 54 honest responders in Rogers's study and our 23 post-MVA patients, is evidence of poor specificity of the SIMS.

After examining the SIMS total scores of their psychiatric patients, Rogers's team concluded "*With this inpatient sample, the recommended SIMS total cut score of >14 functioned poorly, because most genuine responders also exceeded this cut score (i.e., specificity = .28)*" ..... "*more than two-thirds (positive predictive power, PPP = .70) of those identified will be genuine responders. At least for these inpatients, a much higher SIMS total cut score (> 44) is required to achieve a very high specificity.*"[2]

The need to increase the cutoff of SIMS total score to > 44 in the group of legitimate psychiatric patients is consistent with the finding in other studies that the Psychosis (P), Affective Disorders (AF), Neurologic Impairment (NI), and Amnestic Disorders (AM) scales of the SIMS list legitimate medical symptoms[4,5,6] potentially endorsed by both patients and malingerers at similar rates and that SIMS Low Intelligence scale

(LI) consists mainly of arithmetic and logical reasoning tasks or tasks assessing general knowledge[7] on which patients debilitated by a chronic physical or psychiatric illness, or those with the post-concussion syndrome, or persons whose attentional focus is disrupted by chronic pain or psychiatric symptoms may perform worse than uninjured persons and similarly to malingerers.[7,8]

Table 1. *Means and SDs of 23 patients injured in high impact MVAs and of patients in Rogers's 2014 SIMS study*

ANOVA Groups:	N	SIMS total	NI	AM	LI	AF	P	RS	SC
Patients injured in high impact MVAs - data from Cernovsky et al.[5] (2019)	23	26.5 (16.0)	5.2 (3.9)	5.0 (4.4)	4.8 (4.4)	7.7 (2.2)	3.7 (5.3)	4.6 (3.7)	2.7 (2.0)
Mental patients instructed to respond honestly (Rogers et al, 2014)[2]	54	22.6 (11.4)	4.8 (3.4)	5.9 (3.7)	2.0 (1.9)	7.3 (2.1)	2.6 (2.8)	0.7 (2.0)	2.3 (1.9)
Mental patients instructed to exaggerate their symptoms (Rogers et al, 2014)[2]	53	45.0 (14.7)	10.2 (3.6)	10.6 (4.1)	6.0 (3.7)	9.3 (2.1)	9.1 (4.8)	6.1 (4.2)	7.4 (3.0)
Cutoff scores as stipulated by the SIMS manual[3] for its traditional 6 scales (see the next 6 columns) and those determined for RS and SC scales by Rogers et al.[2] (see last 2 columns)		>14	>2	>2	>2	>5	>1	>6	>6

Legend: *NI=SIMS Neurologic Impairment scale, AM=SIMS Amnestic Disorders scale, LI=SIMS Low Intelligence scale, AF=SIMS Affective Disorders scale, P=SIMS Psychosis scale. The mean score of Rogers's 54 honestly responding patients on LI scale was 2.02 prior to rounding, i.e., above the cutoff of 2 points.*

Most of the 23 patients (82.6%) injured in high impact MVAs obtained RS scores below the cutoff specified by Rogers at al. Only 4 patients (17.4%) obtained higher scores, i.e., within the malingering category.

Only one patient (4.3%) scored above 6 points on the SC scale: the rest (95.7%) performed within the non-malingering category. These proportions of probable true negatives on RS and SC scales seem within acceptable limits and comparable to those of some other reputable psychological measures.

Score Differences in ANOVAs

Differences between the 3 groups (those listed in Table 1) were examined in ANOVAs calculated separately for each of the usual 6 SIMS scales and also for Rogers's RS and SC scales. All ANOVAs were statistically significant. The results of post hoc tests are listed in Table 2.

Results with respect to Rogers's SC scale confirm to expectations: the scale performed well. The 54 mental patients responding honestly have not differed significantly from patients injured in high impact MVAs. Both groups obtained significantly lower SC

scores than the 53 mental patients instructed to exaggerate their medical symptoms.

On the RS scale, the patients injured in high impact MVAs did not differ significantly from mental patients instructed to exaggerate their symptoms and these both groups scored significantly higher than mental patients responding honestly. This is a damaging statistical evidence against the RS scale. From a clinical perspective, an inspection of the actual content of RS scale showed that 5 of its 15 items (see RS items with SIMS numbers 56, 58, 67, 73, and 75) are algebraic or logical reasoning tasks on which patients with post-concussion syndrome, pain related insomnia, and distractions caused by pain (i.e., patients with typical post-MVA symptoms) might perform worse than psychiatric patients with only PTSD or mood disorders. Furthermore, our MVA patients were all Canadians (unlike the Rogers's 54 and 53 psychiatric patients) and two of the RS items deal with US history and US geography (RS items with SIMS numbers 14, 63). In Canada, some high school graduates never had school classes in US history or US geography and some failed to acquire that knowledge on their own.

Table 2. *Post hoc tests in ANOVAs on SIMS scales and Rogers's RS and SC scales*

RS scale	***F=(2,127)=35.9, p<.0001*** Tukey HSD Post-hoc Tests: high impact MVA vs genuine MHD: Diff=-3.9000, 95%CI=-5.8816 to -1.9184, p=0.0000 high impact MVA vs exaggerating MHD: Diff=1.5000, 95%CI=-0.4872 to 3.4872, p=0.1770 genuine MHD vs exaggerating MHD: Diff=5.4000, 95%CI=3.8611 to 6.9389, p=0.0000
SC scale	***F=(2,127)=66.4, p<.0001*** Tukey HSD Post-hoc Tests: high impact MVA vs genuine MHD: Diff=-0.4000, 95%CI=-1.8323 to 1.0323, p=0.7858 high impact MVA vs exaggerating MHD: Diff=4.7000, 95%CI=3.2636 to 6.1364, p=0.0000 genuine MHD vs exaggerating MHD: Diff=5.1000, 95%CI=3.9877 to 6.2123, p=0.0000
Total SIMS	***F=(2,127)=38.5, p<.0001*** Tukey HSD Post-hoc Tests: high impact MVA vs genuine MHD: Diff=-3.9000, 95%CI=-11.9757 to 4.1757, p=0.4882 high impact MVA vs exaggerating MHD: Diff=18.5000, 95%CI=10.4016 to 26.5984, p=0.0000 genuine MHD vs exaggerating MHD: Diff=22.4000, 95%CI=16.1288 to 28.6712, p=0.0000
NI	***F=(2,127)=34.4, p<.0001*** Tukey HSD Post-hoc Tests: high impact MVA vs genuine MHD: Diff=-0.4000, 95%CI=-2.5096 to 1.7096, p=0.8947 high impact MVA vs exaggerating MHD: Diff=5.0000, 95%CI=2.8844 to 7.1156, p=0.0000 genuine MHD vs exaggerating MHD: Diff=5.4000, 95%CI=3.7618 to 7.0382, p=0.0000
AM	***F=(2,127)=24.7, p<.0001*** Tukey HSD Post-hoc Tests: high impact MVA vs genuine MHD: Diff=0.9000, 95%CI=-1.4581 to 3.2581, p=0.6381 high impact MVA vs exaggerating MHD: Diff=5.6000, 95%CI=3.2353 to 7.9647, p=0.0000 genuine MHD vs exaggerating MHD: Diff=4.7000, 95%CI=2.8688 to 6.5312, p=0.0000
LI	***F=(2,127)=21.1, p<.0001*** Tukey HSD Post-hoc Tests: high impact MVA vs genuine MHD: Diff=-2.8000, 95%CI=-4.7102 to -0.8898, p=0.0020 high impact MVA vs exaggerating MHD: Diff=1.2000, 95%CI=-0.7155 to 3.1155, p=0.3012 genuine MHD vs exaggerating MHD: Diff=4.0000, 95%CI=2.5167 to 5.4833, p=0.0000
AF	***F=(2,127)=12.7, p<.0001*** Tukey HSD Post-hoc Tests: high impact MVA vs genuine MHD: Diff=-0.4000, 95%CI=-1.6504 to 0.8504, p=0.7290 high impact MVA vs exaggerating MHD: Diff=1.6000, 95%CI=0.3461 to 2.8539, p=0.0084 genuine MHD vs exaggerating MHD: Diff=2.0000, 95%CI=1.0290 to 2.9710, p=0.0000
P	***F=(2,127)=34.6, p<.0001*** Tukey HSD Post-hoc Tests: high impact MVA vs genuine MHD: Diff=-1.1000, 95%CI=-3.5751 to 1.3751, p=0.5444 high impact MVA vs exaggerating MHD: Diff=5.4000, 95%CI=2.9179 to 7.8821, p=0.0000 genuine MHD vs exaggerating MHD: Diff=6.5000, 95%CI=4.5779 to 8.4221, p=0.0000

Legend: *MHD=Mental Health Disorders, NI=SIMS Neurologic Impairment scale, AM=SIMS Amnestic Disorders scale, LI=SIMS Low Intelligence scale, AF=SIMS Affective Disorders scale, P=SIMS Psychosis scale.*

Noteworthy are also ANOVAs comparing scores of our 3 groups on the 6 classical SIMS scales. Patients injured in high impact MVAs have not differed significantly from the 54 honestly responding psychiatric patients on any SIMS scales, with the exception of the Low Intelligence (LI) scale. The MVA patients obtained higher LI scores, perhaps because LI scale contains algebraic or logical reasoning tasks on which patients with post-concussion syndrome, pain related insomnia, and distractions caused by

pain may perform worse than psychiatric patients with only PTSD or mood disorders. In addition, the LI scale also includes the two items from US history and US geography (the MVA patients were all Canadians, unlike the 54 and 53 psychiatric patients).

The 53 psychiatric patients instructed to exaggerate their medical symptoms scored significantly higher on all SIMS scales than the 23 patients injured in high impact MVAs, with the exception of the Low Intelligence (LI) scale on which there was no significant difference between these two groups.

As already mentioned, the LI scale contains logical and algebraic tasks which may pose problems for post-MVA patients due to their post-concussive and other symptoms: this could inflate their LI scores.

Content Analysis of the Rare Symptoms (RS) Scale

The name of the RS scale is a misnomer: its items do not deal with rare symptoms, except perhaps for the one "There are six days in a week" which, as discussed by van Impelen, Mercklebach, Jelic, and Merten,[11] is strictly logically correct and could be perhaps endorsed by some patients with Asperger or psychosis, or also by some intelligent but ornery persons.

It its present form, the RS scale is excessively contaminated by items indicative of thought disorder (RS items with SIMS numbers 3, 8, 28. 42, and 69) thus making it diagnostically unsuitable for that subtype of psychiatric patients and, as already mentioned, also by items involving algebraic or logical reasoning tasks (RS items with SIMS numbers 56, 58, 67, 73, and 75) on which patients with low intelligence, or those with post-concussive symptoms, and/or pain related insomnia, and/or simply distractions caused by persistent pain (e.g., war veterans) could potentially accumulate points towards being misclassified as "malingerers." In this manner, many earlier criticisms of the SIMS LI scale[7,8] and SIMS P scale[6] also now apply to the RS scale.

Concerns About the Symptom Combination (SC) Scale

The SC scale consists of 13 pairs of symptoms, all of these adopted from the SIMS. As already explained, Rogers's team selected those SIMS items for their SC scale which *"were uncorrelated or negatively correlated for genuine responders …"* but *"positively correlated for feigners and accounted for more than 10% of the variance (φ coefficient >.35)."* [2]

Unfortunately, Rogers's psychiatric sample was diagnostically mixed: this makes generalizations of intercorrelations of items from their sample to other diagnostic groups uncertain and highly problematic when scrutinized along APA standards for development of new tests.[13]

When developing future new SC scales, more extensive research would be needed to demonstrate if, or when, specific pairs of symptoms are indeed always mutually exclusive or uncorrelated when separately examined in various well-defined diagnostic groups.

Discussion

The innovative approach of Rogers's team in the development of the Rare Symptoms (RS) scale is greatly superior to previous investigations of the SIMS, but should be undertaken on non-SIMS items in order to avoid listing common and legitimate medical symptoms. The name of this scale is indeed a misnomer.

Future attempts at development of RS and SC scales need to focus on choosing truly rare symptoms for the RS scale and more generally mutually exclusive medical symptoms for the SC scale, and should also include validations on strictly homogeneous diagnostic criterion groups. The mixed sample of psychiatric patients on which the SC and RS scales were developed was not diagnostically adequately homogeneous to meet APA requirements for test validation. The diagnosis should be limited to only one well defined psychiatric condition, separately for each calculation of criterion validity.

There are too many reasons to support a moratorium on clinical, forensic, or other "diagnostic" uses of SIMS by psychologists. Among other various other dire SIMS flaws, even normative SIMS data of healthy college students, listed by the SIMS manual,[3] show that their mean scores on SIMS Psychosis (P) scale (0.8, SD=1.0), Low Intelligence (LI) scale (1.4, SD=1.2), and on Amnestic Disorders (AM) scale (1.2, SD=1.5) are too close to cutoffs for these scales stipulated by SIMS manual. If using the assumption of normal distribution, the z score position of these cutoffs stipulated by SIMS manual (> 1 for P scale, > 2 for LI scale, and also > 2 for AM scale)[3] would absurdly suggest that perhaps 42.1% of these college students malinger psychotic

symptoms, 30.9% low intelligence, and 29.8% the amnestic disorder, see recent statistical re-evaluation of Glenn Smith's normative data by Cernovsky and Fatahi.[10]

This irremediable flaw of Smith's normative "normal" samples has an equivalent in misclassifying legitimate patients as malingerers: the mean scores on all 6 SIMS scales of Rogers's 54 psychiatric patients responding honestly are in the category of malingering (see Table 1; as already explained, even their LI score was 2.02 prior to rounding to one decimal point, i.e., above the cutoff of > 2).[2] Similarly, all SIMS scores of the 23 high impact MVA patients were above the cutoffs on the 6 classical SIMS scales, see Table 1.

It has been demonstrated that more than 50% of SIMS items reflect post-MVA symptoms.[20] By honestly reporting these symptoms, legitimate patients paradoxically receive one point for each of their symptoms towards the diagnosis of "malingering." With respect to survivors of car accidents, recent meta-analytic ANOVAs[9] showed that lowest SIMS scores were obtained by meta-analytically combined nine samples of normal controls (combined N=500), the next higher scores were those of 47 patients with minor whiplash symptoms from car accidents, and the next higher scoring group were 23 patients more severely injured in high impact MVAs: these 3 groups differed significantly in the number of reported typical post-MVA symptoms, i.e., thus in their SIMS scores. These ANOVAs also included a meta-analytically combined sample of 88 persons instructed to feign post-MVA symptoms (whiplash, post-concussive symptoms, or both): SIMS scores of these instructed malingerers were usually statistically similar to those of patients injured in high impact MVAs.[9] This shows that the SIMS differentiates between reporters and non-reporters of medical symptoms, but not malingerers from injured patients.

Since post-MVA symptoms are often similar to polytraumatic symptom patterns encountered in US and Canadian war veterans, that group of patients is also likely to feel defrauded by SIMS psychologists. Even supposedly cautious statements by SIMS psychologists such as "*this patient may be exaggerating, magnifying, or over-reporting his symptoms*" trigger the unethical professional consequence of making the patient suspect to insurance clerks, and thus result in denials or undue delays of therapies and of other legally owed benefits.

The readers are also cautioned against a blind use of the SIMS based RS and SC scales. Some investigators complained about low sensitivity of these scales, i.e., about their low capacity to identify malingerers. This could indeed perhaps happen if malingerers report symptoms not listed in the RS scale, or not adequately represented in that scale. For example, a recent data analysis of RS and SC scales by Edens, Truong, and Otto[21] reported that RS and SC cut-off scores determined by Rogers's team "*to classify individuals as faking did not work particularly well, suggesting that further research is needed before these scales can be used in applied settings.*"

Eden's team complained about the low sensitivity of the published cutoffs of > 6 for the RS and SC scales, and they experimented with lower cutoffs which they found more satisfactory "*Lower potential cut-offs did, however, improve sensitivity to feigning somewhat while not excessively diminishing specificity*."

In our sample of 23 survivors of high impact car accidents in which their vehicle was damaged to the extent of being deemed not worthy of repair, lower cutoffs would likely misclassify the following proportions as "malingerers." In our sample of MVA patients, as already mentioned, the cutoff of > 6 classified 17.4% as malingerers via RS and 4.3% via SC scale. No patient obtained the score of 6 on the RS, but two patients did on the SC scale: thus, the cutoff > 5, classified also only 17.4% as malingerers via RS, but 8.7% via SC; the cutoff >4, classified so 30.4% by RS and 26.1% by SC; the cutoff > 3 classified so 39.1% by RS and 26.1% by SC (there was no patient with the score of 4 on SC scale, so the proportions were the same for cutoffs > 4 and > 3). In general, maximizing the proportions of detected malingerers at the cost of more frequently infringing on human rights and dignity of those denied therapy for being misclassified as malingerers appears incongruent with the medical goal of *beneficence, not maleficence*, sometimes formulated as *primum non nocere.* For instance, an insurance contracted psychologist with office assistants might assess 300 patients over the next 2 years and thus misdiagnose, via cutoff >5, about 52 legitimate patients as malingerers via RS scale, or 91 persons with RS cutoff > 4. The reader needs to consider the content of RS items, especially the 5 tasks involving logical reasoning or algebraic tasks. The extreme sudden acceleration or deceleration in motor vehicle collisions causes microvascular injuries

and axonal shearing with subsequent neurotoxicity. Such injuries would interfere with performance on algebraic or logical reasoning tasks of the RS scale and presumably also with the inter-item consistency of responses on the SC scale. In a surprisingly absurd manner, those more severely injured would obtain higher RS and SC scores, thus being more likely misclassified as "malingerers." This is why, in the future, SIMS style items should not be included in constructing "rare symptoms" (RS) scales, and similarly, the clinical usefulness of SIMS style SC items needs to be reconsidered as they would also need to be validated separately on each diagnostically more homogeneous group to comply with APA standards.

Given the content of RS and SC scales currently proposed by Rogers, patients at the highest risk of being harmed by misdiagnosis of malingering are probably not only the survivors of high impact MVAs, but also war veterans with post-concussive symptoms and insomnia caused by chronic pain, or persons with similar polytraumatic symptom patterns from industrial accidents, or also incarcerated patients with psychotic symptoms in forensic settings. The more intense their symptoms, the higher would be their RS and SC scores.

Furthermore, the innovative approaches such as by Rogers's team in creating RS and SC scales are also methodologically complicated by a conceptual distinction between "*symptom accentuators*" (patients overreporting their existing symptoms) and "*symptom producers*" (malingerers entirely free of symptoms), as recently described by Mazza's group[22]. Rogers's study[2] compared honestly responding psychiatric patients only to those exaggerating their symptoms, i.e., to symptom accentuators, not to symptom producers and this probably affected the item composition of RS and SC scales.

To differentiate between such groups, Mazza's team evaluated persons with diagnosed adjustment disorder with mixed anxiety and depressed mood who had undergone assessment for psychiatric/psychological damage: 39 were classified as producers, 44 as accentuators, and 49 as regular patients who neither exaggerated nor feigned their symptoms. SIMS scores of these three groups were compared. The results showed that SIMS scales that appeared to best differentiate the groups were the SIMS total score, Neurologic Impairment (NI) scores, and Low

Intelligence (LI) scores. Obviously, in this particular sample of patients with preliminary diagnosis of only adjustment disorder with mixed anxiety and depression, both the regular patients and the accentuators were less likely to stray into endorsing SIMS neuropsychological items and also to score in the "malingering direction" on logical or algebraic reasoning tasks of the LI scale than the producers. As shown in another investigation,[20] more than 50% of SIMS items overlap conceptually with symptoms listed in scales measuring post-concussion (Rivermead scale[17]) and subjective symptoms of whiplash (PMNS[16]). Endorsing such neuropsychological SIMS items inflated the SIMS total scores of the "producers" in Mazza's study. Patients with only depression and anxiety appropriately reported or only over-emphasized their existing symptoms without also indiscriminately feigning neuropsychological symptoms of the NI scale[5] and of the LI scale[7] as did the "symptom producers."

It should be noted in this context that some MVA patients also most probably exaggerate the scope of their symptoms, perhaps on the pragmatic and not quite unreasonable assumption that their legitimate medical complaints could otherwise remain ignored by clerks of the car insurance company. While waiting somewhere in physicians' waiting rooms for assessments, post-MVA patients might overhear worrisome rumors of high rejection rates, by car insurance clerks, of legitimate post-MVA claims by injured persons. Such unjust rejections indeed happen frequently because many insurance contracted psychologists still use the SIMS and/or Paul Green's tests[23] to falsely declare injured persons as malingerers. A study based on Toronto's Medex files (mentioned in a publication on Gutierrez questionnaire[24], see page 16), found the rejection rate in May 2019 amounting to 73.7% of applicants for post-MVA medical therapies and other insurance benefits. Such facts make chastising of accentuators for "*overreporting or magnifying their symptoms*" somewhat inappropriate and perhaps even somewhat reprehensible. The same may be true about accentuators within other patient groups such as injured war veterans or patients with industrial injuries.

The most infamous test that has allowed countless false rejections of such legitimate patients is the SIMS: it is favored by many insurance contracted psychologists because it does not require much

administration time and scoring time. Statistical meta-analyses have demonstrated conclusively that the SIMS differentiates only between reporters and non-reporters of legitimate medical symptoms, but not between legitimate patients and malingerers.[9] As already explained, patients with more intense or extensive symptoms are more likely to be classified as malingerers by the traditional SIMS scales than less injured patients and the same is potentially true also for the new (but SIMS based) RS and SC scales.

Another serious methodological concern about Rogers's RS and SC scales is that, no matter what item pool is used to statistically extract items for these scales with the intelligent and innovative procedure developed by Rogers, the type and content of the extracted items might differ very widely depending on what diagnostic groups are used in the procedure. The APA standards would require a separate validation for each diagnostic dimension.

Conclusions

Rogers's RS and SC scale may appear to perform adequately on some samples, but generalizability of their cutoffs to patients with other psychiatric diagnosis and especially to other medical groups remains too uncertain. The RS scale is contaminated by its inclusion of SIMS items consisting of algebraic or logical reasoning tasks and of those representing delusional ideation or thought disorder: in an absurd manner, more injured or more severely ill patients are likely to be misclassified more frequently as malingerers than those less inured or less ill. The innovative approach by Rogers in his development of RS and SC scales should be now applied to items other than those from the SIMS, with the focus on choosing truly rare symptoms for the RS scale, and on selecting more generally mutually exclusive medical symptoms for the SC scale, and should also include validations on more narrowly defined diagnostic criterion groups, separately for each specific diagnostic category.

References

[1] Rogers R and Bender SD (editors). *Clinical Assessment of Malingering and Deception*. 4th Edition. New York, NY: The Guilford Press, Inc., 2018.

[2] Rogers R, Robinson EV, and Gillard ND. The SIMS Screen for Feigned Mental Disorders: the Development of Detection-based Scales. *Behavioral Sciences and the Law*. 2014; 32: 455–466.

[3] Widows MR, and Smith GP. *Structured Inventory of Malingered Symptomatology - Professional Manual*. Lutz, FL: PAR Inc., 2005.

[4] Cernovsky ZZ, Mendonça JD, Ferrari JR, Sidhu G, Velamoor V, Mann SC, Oyewumi LK, Persad E, Campbell R, and Woodbury-Fariña MA. Content Validity of the Affective Disorder Subscale of the SIMS. *Archives of Psychiatry and Behavioral Sciences*. 2019;2(2):33-39.

[5] Cernovsky Z, Bureau Y, Mendonça J, Velamoor V, Mann S, Sidhu G, Diamond DM, Campbell R, Persad E, Oyewumi LK, and Woodbury-Fariña MA. Validity of the SIMS Scales of Neurologic Impairment and Amnestic Disorder. *International Journal of Psychiatry Sciences*. 2019; 1(1):13-19.

[6] Cernovsky Z, Mendonça JD, Oyewumi LK, Ferrari JR, Sidhu G, and Campbell R. Content Validity of the Psychosis Subscale of the Structured Inventory of Malingered Symptomatology (SIMS). *International Journal of Psychology and Cognitive Science*. 2019; 5(3): 121-127.

[7] Cernovsky ZZ, Mendonça JD, Ferrari JR, Bureau YRJ. Content validity of SIMS low intelligence scale. *International Journal of Research in Medical Science*. 2019;1(1):21-25.

[8] Cernovsky ZZ, Mendonça JD, Bureau YRJ, and Ferrari JR. Criterion Validity of Low Intelligence Scale of the SIMS. *International Journal of Psychology Sciences*. 2019;1(1):3-5.

[9] Cernovsky ZZ, Mendonça JD, Ferrari JR. Meta-Analysis of SIMS Scores of Survivors of Car Accidents and of Instructed Malingerers. *Archives of Psychiatry and Behavioral Sciences*. 2020; 3(1): 01-11.

[10] Cernovsky Z and Fattahi M. Meta-analysis of Scale Cutoffs in the Structured Inventory of Malingered Symptomatology (SIMS). *International Journal of Psychology Sciences*. 2020;1(1):1-7.

[11] Van Impelen A, Merckelbach H, Jelicic M, Merten T. The Structured Inventory of Malingered Symptomatology (SIMS): a systematic review and meta-analysis. *The Clinical Neuropsychologist*. 2014; 28(8): 1336-1365. doi: 10.1080/13854046.2014.984763

[12] Clegg C, Fremouw W, and Mogge NL. Utility of the Structured Inventory of Malingered Symptomatology (SIMS) and the Assessment

of Depression Inventory (ADI) in screening for malingering among outpatients seeking to claim disability. *Journal of Forensic Psychiatry and Psychology*. 2009: 20(2): 239-254. doi: 10.1080/14789940802267760

[13] American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. *The Standards for Educational and Psychological Testing*. Washington, D.C.: AERA Publications, 2014.

[14] Cernovsky ZZ, Istasy PVF, Hernández-Aguilar ME, Mateos-Moreno A, Bureau Y, and Chiu S. Quantifying Post-Accident Neurological Symptoms Other than Concussion. *Archives of Psychiatry and Behavioral Sciences*. 2019; 2(1): 50-54.

[15] Cleeland CS. *The Brief Pain Inventory - User Guide*. Houston, TX: The University of Texas - M. D. Anderson Cancer Center, 2009.

[16] Morin CM, Belleville G, Bélanger L, and Ivers H. The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep*. 2011; 34:601-608.

[17] Eyres S, Carey A, Gilworth G, Neumann V, Tennant A. Construct validity and reliability of the Rivermead Post-Concussion Symptoms Questionnaire. *Clinical Rehabilitation*. 2005; 19: 878-87.

[18] Omalu BI, DeKosky ST, Minster RL, Kamboh MI, Hamilton RL, & Wecht CH. Chronic traumatic encephalopathy in a National Football League player. *Neurosurgery*. 2005; 57: 128-34.

[19] Omalu BI, DeKosky ST, Hamilton RL, Minster RL, Kamboh MI, Shakir AM, & Wecht CH. Chronic

traumatic encephalopathy in a National Football League player: Part II. *Neurosurgery*. 2006; 59: 1086-92.

[20] Cernovsky ZZ, Ferrari JJR, Mendonça JD. Pseudo -diagnoses of Malingering of Neuropsychological Symptoms in Survivors of Car Accidents by the Structured Inventory of Malingered Symptomatology. *Archives of Psychiatry and Behavioral Sciences*. 2019; 2(1): 55-65.

[21] Edens JF, Truong TN, Otto RK. Classification accuracy of the rare symptoms and symptom combinations scales of the Structured Inventory of Malingered Symptomatology in three archival samples. *Law and Human Behavior*. 2020 Jan 23. doi: 10.1037/lhb0000361

[22] Mazza C, Orru` G, Burla F, Monaro M, Ferracuti S, Colasanti M, Roma P. Indicators to distinguish symptom accentuators from symptom producers in individuals with a diagnosed adjustment disorder: A pilot study on inconsistency subtypes using SIMS and MMPI-2-RF. *PLoS ONE*. 2019; 14(12): e0227113. https://doi.org/10.1371/journal.pone.0227113

[23] Ferrari JJR, Cernovsky ZZ, Mendonça JD. False Positives in Green's Tests of Malingering on Chronic Pain Patients. *International Journal of Psychology and Cognitive Science*. 2019; 5(2): 58-63.

[24] Gutierrez J, Nosonova V, Cernovsky Z, Fattahi M, Tenenbaum S. Gutierrez Questionnaire for Assessments of Patients after Car Accidents. *Archives of Psychiatry and Behavioral Sciences*. 2019; 2(2):10-21.