A. Purpose

This AM provides guidance to all State and Federal adjudicators (including administrative law judges) on how to consider Global Assessment of Functioning (GAF) ratings when assessing disability claims involving mental disorders. GAF ratings are frequently included in medical evidence and consultative examination reports from mental health care professionals. We consider a GAF rating as opinion evidence. As with other opinion evidence, the extent to which an adjudicator can rely on the GAF rating as a measure of impairment severity and mental functioning depends on whether the GAF rating is consistent with other evidence, how familiar the rater is with the claimant, and the rater’s expertise.

B. Background

The Diagnostic and Statistical Manual of Mental Disorders (DSM), published periodically by the American Psychiatric Association (APA), provides the common language and standard criteria for classification of mental disorders. The DSM, Fourth Edition, Text Revision (DSM-IV-TR) provided for a multi-axial assessment of mental disorders. Axis I covered clinical disorders; Axis II covered personality disorders and mental retardation; Axis III covered general medical conditions; Axis IV considered psychosocial and environmental problems; and Axis V, also known as the GAF, reported the clinician’s judgment of a person’s overall level of functioning.
NOTE: Recently the APA published a fifth edition (DSM-5) that does not include GAF rating for assessment of mental disorders. However, we will continue to receive and consider GAF in medical evidence.

C. What is the GAF?

The GAF is a rating reporting the clinician’s judgment of a person’s ability to function in daily life. See DSM-IV-TR, p. 32. It reflects the clinician’s subjective judgment about the person’s symptom severity and psychological, social, and occupational functioning. The rating does not reflect impairment in function caused by physical or environmental limitations.

Each 10-point range within the GAF has two components: one that covers symptom severity, and a second covering functioning. If a person’s symptom severity and level of functioning differ, the GAF rating reflects the lower rating. In other words, a person with a number of psychological symptoms and very few functional limitations would get a GAF score consistent with his or her reported psychological symptoms (DSM-IV-TR, p. 33).

D. Problems with using the GAF to evaluate disability

The problem with using the GAF to evaluate disability is that there is no way to standardize measurement and evaluation. Researchers have pointed out numerous problems with the GAF - from the choice of rating scale, to the choice of anchor points, to the limited instructions to clinicians for rating GAFs within the 10-point intervals. Some researchers suggested that abandoning the GAF altogether in favor of a 5-point scale, might make ratings more accurate.

Other researchers had questions about the anchor points of the scale, because it is unclear whether other key words or examples would result in a better GAF. The anchor points are indicators that help the clinician establish the GAF rating. The current anchor points are generally inclusion criteria. For example, “occasionally failing to maintain personal hygiene” is a criterion that could result in a GAF rating of 20-11. Some researchers believe adding exclusion criteria to the anchors might promote better inter-rater reliability. An exclusion criterion might be “do not rate GAF lower than
20 if the individual is clean but somewhat disheveled”.

As another example, some researchers have expressed concern that, as managed-care companies continue to set practice standards and quality guidelines, and the trend towards accountability for clinical outcomes continues to grow, practitioners may be under increasing pressure to use GAF ratings and DSM diagnoses in inappropriate ways to ensure insurance coverage for treatment.

Research has also identified the propensity of some clinicians to give inflated or unrealistically low GAF ratings because the GAF rating instructions in the DSM-IV-TR are unclear. In general, inter-rater reliability ratings are low in the clinical setting because there is great variability of training and experience levels amongst clinicians. These rating problems, alone or in combination, can lead to improper assessment of impairment severity.

1. **GAF ratings are not standardized**

The GAF is neither standardized nor based on normative data. A GAF rating compares the patient with the distinctive population of patients the clinician has known. This limits direct comparability of GAF scores assigned by different evaluators or even by the same evaluator at substantially different points in time.

Although the GAF rating is numerical, the actual number assigned can be misleading because the score does not quantify differences in function between people. For example, a GAF score of 75 does not mean a person is functioning 10 units better than a person with a score of 65, nor does a GAF of 40 indicate a person is functioning half as well as a person with a score of 80.

GAF ratings assigned by different clinicians are inconsistent because of the lack of standardization. In a clinical (i.e., treatment and evaluation) setting, inter-rater reliability is generally low because of variability of training and experience levels among clinicians. This means adjudicators cannot draw reliable inferences from differences in GAF ratings assigned by different clinicians or from a single GAF score in isolation.
2. **GAF is not designed to predict outcome**

Clinicians use the GAF to help plan and measure the impact of treatment (see DSM-IV-TR, p. 32). It is of limited value for assessing prognosis or treatment outcome and other measures are better indicators of outcome. For example, research suggests the highest level of functioning for a period of time is a better predictor of prognosis than the lowest level of functioning for a time period.

3. **GAF ratings need supporting detail**

The GAF scale anchors are very general and there can be a significant variation in how clinicians rate a GAF. For example, if a claimant has a GAF of 20, it could mean that he or she is not maintaining minimal personal hygiene (a clinical observation), or that the claimant has some potential to hurt himself or others (a clinical judgment). Interpreting the GAF rating requires knowing what the clinician was focusing on when assigning the overall rating. Although the DSM-IV-TR recommends assigning a rating reflecting the lower assessment when symptom severity and level of functioning differ, evaluators rarely note whether the score reflects function, symptoms, or both.

**E. Weighing GAF ratings as opinion evidence**

For purposes of the Social Security disability programs, when it comes from an acceptable medical source, a GAF rating is a medical opinion as defined in 20 CFR §§ 404.1527(a)(2) and 416.927(a)(2). An adjudicator considers a GAF score with all of the relevant evidence in the case file and weighs a GAF rating as required by §§ 20 CFR 404.1527(c), 416.927(c), and SSR 06-03p, while keeping the following in mind:

The GAF is unlike most other opinion evidence we evaluate because it is a rating. However, as with other opinion evidence, a GAF needs supporting evidence to be given much weight. By itself, the GAF cannot be used to “raise” or “lower” someone’s level of function. The GAF is only a snapshot opinion about the level of functioning. It is one opinion that we consider with all the evidence about a person’s functioning. Unless the clinician
clearly explains the reasons behind his or her GAF rating, and the period to which the rating applies, it does not provide a reliable longitudinal picture of the claimant’s mental functioning for a disability analysis.

A GAF score is never dispositive of impairment severity. **DO NOT:**

Give controlling weight to a GAF from a treating source unless it is well supported and not inconsistent with the other evidence.

When case evidence includes a GAF from a treating source and you do not give it controlling weight, you must provide good reasons in the personalized disability explanation or decision notice.

Equate any particular GAF score with a listing-level limitation.

You cannot use a GAF rating to determine whether a claimant’s impairment meets the diagnostic criteria of mental retardation in listing 12.05, because the rating lacks specificity, may not reflect a claimant’s functioning over time, and is not a standardized measure of anything, including intelligence and adaptive behavior.

Equate a particular GAF score with a particular mental residual functional capacity assessment.

The GAF does not measure the ability to meet the mental demands of unskilled work. There have been no published studies of how, or if, GAF scores relate to meeting the demands of unskilled work. Additionally, there is no correlation between GAF scores and the B criteria in the mental disorders listings.

**Questions**

Direct all program–related and technical questions to your RO support staff or PSC OA staff. RO support staff or PSC OA staff may refer questions or problems to their Central Office contacts.

**References:**

20 CFR 404.1502 General definitions and terms for subpart P
20 CFR 416.902 General definitions and terms for subpart I
20 CFR 404.1520a Evaluation of mental impairments
20 CFR 416.920a Evaluation of mental impairments
20 CFR 404.1527 Evaluating opinion evidence
Evaluating opinion evidence

SSR 96-2p Titles II and XVI: Giving controlling Weight to Treating Source Medical Opinions

SSR 06-03p Titles II and XVI: Considering Opinions and Other Evidence from Sources Who Are Not “Acceptable Medical Sources” in Disability Claims; Considering Decisions on Disability by Other Governmental and Nongovernmental Agencies

DI 22505.001 Medical Evidence of Record (MER) Policies

DI 22505.003 Medical and Other Evidence of an Individual’s Impairment(s)

DI 22505.007 Developing Initial Evidence from Medical Sources

DI 24510.065 Section III of SSA-4734-F4-SUP Functional Capacity Assessment

DI 24515.002 Evaluating Opinion Evidence – Basic Policy

DI 24515.003 Weighing Medical Opinions from Treating Sources and Other Medical Sources

DI 24515.004 Giving Controlling Weight to Treating Source Medical Opinions (SSR 96-2p)

DI 24515.005 Evaluating Noncontrolling Treating Source Medical Opinions

DI 24515.006 Evaluating Nontreating Source Medical Opinions

DI 24515.008 Titles II and XVI: Considering Opinions and Other Evidence from Sources Who Are Not “Acceptable Medical Sources” in Disability Claims; Considering Decisions on Disability by Other Governmental and Nongovernmental Agencies (SSR 06-03p)

DI 26530.015 Personalized Disability Explanation in Initial Closed Period, Unfavorable Onset Date, and DWB Allowances Where the Month Of Entitlement is Restricted to January 1991 Under P.L. 101-508

DI 26530.020 Personalized Disability Explanation in Initial Denials

DI 33015.020 Writing the Disability Hearing Officer’s (DHO’s) Decision

HALLEX 1-2-5-1 Evidence – General

HALLEX 1-2-5-14 Obtaining Medical Evidence from a Treating Source or Other Medical Source


Tineke van Veen MD PhD, Nic J. A. van der Wee MD PhD, Frans G. Zitman MD PhD.
Global Assessment of Functioning (GAF): properties and frontier of current knowledge. Aas IH.