DIAGNOSTIC AND PROGNOSTIC SIGNIFICANCE OF DSM-5 ATTENUATED PSYCHOSIS SYNDROME IN SERVICES FOR INDIVIDUALS AT ULTRA HIGH RISK FOR PSYCHOSIS

-Invited manuscript for Schizophrenia Bulletin Special Issue-

Fusar-Poli Paolo\textsuperscript{1,2,3}, De Micheli Andrea\textsuperscript{1,3}, Cappucciati Marco\textsuperscript{1,3}, Rutigliano Grazia\textsuperscript{1,3}, Davies Cathy\textsuperscript{1,3}, Ramella-Cravaro Valentina\textsuperscript{1,3}, Oliver Dominic\textsuperscript{1,3}, Bonoldi Ilaria\textsuperscript{1,2,3}, Rocchetti Matteo\textsuperscript{1,3}, Gavaghan Lauren\textsuperscript{2}, Rashmi Patel\textsuperscript{1,2}, McGuire Philip\textsuperscript{1,2}

Affiliations

1. Department of Psychosis Studies, King's College London, Institute of Psychiatry Psychology and Neuroscience, London, United Kingdom;

2. OASIS service, South London and the Maudsley NHS Foundation Trust, London, United Kingdom;

3. Early Psychosis: Interventions & Clinical-detection (EPIC) lab, King's College London, Department of Psychosis Studies, Institute of Psychiatry Psychology and Neuroscience, London, United Kingdom.

Endnote_Neuroprevent

Word count

Abstract: 250

Text: 4000

Figures: 3

Tables: 4

Financial support: This work was supported in part by a 2014 NARSAD Young Investigator Award to Paolo Fusar-Poli.

Conflict of Interests: None
ACRONYMS, ABBREVIATIONS

CAARMS 12/2006, Comprehensive Assessment of the At Risk Mental State 12/2006 version

UHR+, at clinical high risk for psychosis according to the CAARMS 12/2006

UHR-, not at clinical high risk for psychosis according to the CAARMS 12/2006

BLIPS, Brief Limited Intermittent Psychotic Symptoms subgroup of the CAARMS 12/2006

GRD, Genetic Risk and Deterioration Syndrome subgroup of the CAARMS 12/2006

APS, Attenuated Psychotic Symptoms subgroup of the CAARMS 12/2006

DSM-5-APS Attenuated Psychosis Syndrome

DSM-5-APS+, meeting DSM-5-APS+ criteria

DSM-5-APS-, not meeting DSM-5-APS+ criteria

SIPS, Structured Interview for Prodromal Symptoms
ABSTRACT

Background
The diagnostic and prognostic significance of the DSM-5-defined Attenuated Psychosis Syndrome (DSM-5-APS) in individuals undergoing an Ultra High Risk (UHR) clinical assessment for suspicion of psychosis risk is unknown.

Methods
Prospective cohort study including all consecutive help-seeking individuals undergoing both a DSM-5-APS and a Comprehensive Assessment of At Risk Mental State (CAARMS 12/2006) assessment for psychosis risk at the OASIS UHR service (March 2013-April 2014). The diagnostic significance of DSM-5-APS was assessed with percent overall agreement, prevalence bias adjusted kappa, Bowker’s test, Stuart-Maxwell test, residual analysis; the prognostic significance with Cox regression, Kaplan–Meier failure function, time-dependent Area Under the Curve (AUC) and net benefits analysis. The impact of specific revisions of the DSM-5-APS was further tested.

Results
In 203 help-seeking individuals undergoing UHR assessment, the agreement between the DSM-5-APS and the CAARMS 12/2006 was only moderate (kappa 0.59). Among 142 non-psychotic cases, those meeting DSM-5-APS criteria had a fivefold probability (HR 5.379) of developing psychosis compared to those not meeting DSM-5-APS criteria, with a 21-month cumulative risk of psychosis of 28.17% vs 6.49% respectively. The DSM-5-APS prognostic accuracy was acceptable (AUC 0.76 at 24 months) and similar to the CAARMS 12/2006. The DSM-5-APS designation may be clinically useful to guide the provision of indicated interventions within a 7%-35% (two-year) range of psychosis risk. The removal of the criterion E and C of the DSM-
5-APS may improve its prognostic performance and transdiagnostic value.

**Conclusions**

The DSM-5-APS designation may be clinically useful in individuals accessing clinical services for psychosis prevention.

**Keywords:** Psychosis, Schizophrenia, CAARMS, SIPS, Prodromal, DSM-5
‘It is of greatest practical importance to diagnose cases of dementia praecox with certainty and at an early stage’ (Kraeplin, 1896, p.23).

INTRODUCTION

The need to prospectively identify individuals at ultra high risk for psychosis (UHR) stems from the poor outcomes of psychotic disorders under standard care. The Comprehensive Assessment of At Risk Mental States (CAARMS), was developed in Melbourne twenty years ago to assess and rate the UHR criteria of attenuated psychotic symptoms (APS), brief limited intermittent psychotic symptoms (BLIPS) and trait vulnerability plus decline in psychosocial functioning (Genetic Risk and Deterioration Syndrome, GRD). Since then, the CAARMS has been used to predict incipient risk of transition to psychosis -while it does not predict the onset of incident non-psychotic disorders- within a relatively short period of time (2 years). The use of the CAARMS is associated with good prognostic accuracy worldwide, at least in help-seeking individuals recruited on suspicion of psychosis risk who have accumulated several risk factors for the development of psychosis. Over the past two decades, similar instruments have been developed (e.g. the Structured Interview for Prodromal Symptoms, SIPS), each with the goal of facilitating primary indicated interventions in at risk individuals.

More recently, a competing diagnostic construct, the DSM-5 Attenuated Psychosis Syndrome (hereafter DSM-5-APS), has been further proposed to identify attenuated psychotic symptoms. Contrary to popular belief, the DSM-5-APS is not only listed in the research appendix (Section III, starting on page 783) of the DSM-5, but it also appears in the main body of text, in the “Schizophrenia Spectrum and Other
Psychotic Disorders” Section (page 122)\textsuperscript{15}, where it is featured with the official codable diagnosis (298.8) of “Other Specified Schizophrenia Spectrum Disorder and Other Psychotic Disorder”\textsuperscript{15}. Consequently, 120 years after Kraepelin’s claims, two competing diagnostic constructs are available to identify individuals at risk for psychosis: the CAARMS (or SIPS) and DSM-5-APS. Competing diagnostic constructs may be a major source of confusion\textsuperscript{16}. For example, some authors have started using the DSM-5-APS and CAARMS/SIPS designations interchangeably\textsuperscript{17, 18}. However, although both constructs focus on at-risk symptoms, there are profound differences in their operationalization (see Table 1).

*** TABLE 1 ABOUT HERE ***

Initial studies have investigated the basic characteristics of the DSM-5-APS designation in non-clinical samples drawn from schools\textsuperscript{19} or the general population\textsuperscript{20}, and in clinical samples drawn from outpatients\textsuperscript{21}, inpatients\textsuperscript{22}, mixed psychiatric clinics\textsuperscript{23} or professionals working in general practice\textsuperscript{24}. Surprisingly, no study has been designed to directly test the magnitude of diagnostic agreement and disagreement between the DSM-5-APS and the CAARMS within the pool of subjects seeking help at UHR services. Similarly, the prognostic accuracy of the DSM-5-APS designation for the prediction of psychosis is completely unknown.

Here, we present the first study directly investigating the diagnostic and prognostic significance of the DSM-5-APS in help-seeking individuals referred to UHR services and undergoing an assessment on suspicion of psychosis risk. Our first aim was to quantify the diagnostic agreement and disagreement between the DSM-5-APS and the
CAARMS in UHR services. Second, we aimed to estimate the prognostic accuracy of the DSM-5-APS compared to the CAARMS standard and assess its potential clinical usefulness in UHR services. Third, we tested the impact of different DSM-5-APS criteria revisions.

METHODS
Sample

All consecutive individuals referred to the Outreach and Support in South London (OASIS) UHR service on suspicion of psychosis risk in the period March 2013-April 2014, and who underwent a Comprehensive Assessment of At Risk Mental State (CAARMS, version 12/2006)-based UHR assessment25 were included (there were no exclusions). Possible outcomes of the CAARMS 12/2006 assessment are: at risk for psychosis (UHR+), not at risk for psychosis (UHR-) or already psychotic (Psychosis). Details of the OASIS UHR service have been previously described26. The CAARMS 12/2006 is the gold standard assessment instrument at the OASIS. The procedure for staff training, inter-rater reliability measures and the specific types of indicated interventions adopted at OASIS have been detailed in separate manuscripts27, 28. For the purpose of the current study, two psychiatrists underwent an initial 2-hour debriefing on the use of the DSM-5-APS and, under the supervision of the OASIS consultant (who had more than 10 years of experience in the assessment of UHR individuals), were required to additionally score (yes/no) the six DSM-5-APS inclusion criteria (A to F, see Table 1 for details) during the face-to-face CAARMS 12/2006. Importantly, the study sample was designed to reflect at-risk populations as they are encountered in day-to-day practice in UHR services. Avoiding the use of external and non-help-seeking healthy control groups, who do not reflect the clinical
composition of people actually screened in UHR services, is essential to assess the real-world diagnostic and prognostic significance of DSM-5-APS in this population.

**Study measures**

*Primary outcomes*

**Baseline analysis (diagnosis)**

The primary outcome measure for the baseline analysis was the diagnostic agreement and disagreement between the DSM-5-APS (DSM-5-APS+, DSM-5-APS-, Psychosis) and the CAARMS 12/2006 (UHR+, UHR-, Psychosis).

**Longitudinal analysis (prognosis)**

The primary outcome for the longitudinal analysis was the prognostic accuracy of the DSM-5-APS for the prediction of psychotic disorders and its clinical usefulness. Psychosis onset was defined according to the CAARMS 12/2006. The follow-up period began at the time of the baseline CAARMS 12/2006 assessment at the OASIS service, censored at 1st August 2016.

**Secondary measures**

Secondary measures included sociodemographic (age, sex, self-assigned ethnicity) and psychopathological (severity and frequency of UHR symptoms as measured on the P1-P4 subscales of the CAARMS 12/2006) variables, as well as functional level (assessed with the Social and Occupational Functioning Assessment Scale, SOFAS29).

**Statistical analysis**
Cross-sectional analysis (diagnosis)

The diagnostic significance of the DSM-5-APS was addressed by measuring the baseline agreement and disagreement between the DSM-5-APS and the CAARMS 12/2006. The index DSM-5-APS test outcomes (DSM-5APS+, DSM-5-APS-, Psychosis) were thus contrasted against the CAARMS 12/2006 gold standard (i.e. UHR-, UHR+ [APS, BLIPS, GRD], Psychosis). When individuals met criteria for multiple UHR+ subgroups they were classified as follows: any BLIPS>APS or APS+GRD>GRD alone. We estimated percent overall agreement, kappa (with its 95%CI) and the Prevalence And Bias Adjusted Kappa (PABAK), which adjusts the kappa for differences in prevalence and bias (see eMethods1). Since the null hypothesis with kappa is of disagreement, we measured contingency table symmetry with Bowker’s test and marginal homogeneity with the exact Stuart-Maxwell test (these tests are based on a null hypothesis of agreement). Finally, to better elucidate the impact of each specific cell on the overall results, we reported the adjusted residuals and Fisher’s exact test and conducted complementary analyses after excluding the BLIPS and GRD cases.

Longitudinal analysis (prognosis)

We used Cox proportional hazards complete-case analyses to evaluate the effects of the DSM-5-APS designation (DSM-5-APS+ or DSM-5-APS-) on the development of psychotic disorders and time to development of psychosis, after checking the proportional hazards assumption, among those who were not already psychotic at baseline. We also plotted the overall cumulative risk of psychosis onset in individuals assessed for DSM-5-APS with the Kaplan–Meier failure function (1 – survival) and Greenwood 95% CIs. Kaplan-Meier point estimates were also reported, truncated
when at least ten subjects were still at risk. Prognostic accuracy of the DSM-5-APS was further determined with the C statistic. In a first step, we estimated and plotted the time-dependent ROC curve and area under time-dependent ROC curve (AUC) for the DSM-5-APS designation at different timepoints, together with 95% confidence bands of AUCs (values of 0.9-1.0 are considered outstanding, 0.8-0.9 excellent and 0.7-0.8 acceptable\textsuperscript{39}). In a second step, we tested and plotted potential AUCs differences of the two competing designations (DSM-5-APS vs CAARMS 12/2006), measured on the same individuals. Time-dependent Sensitivity (Se), Specificity (Sp), Positive Predictive Values (PPV) and Negative Predictive Values (NPV) were additionally computed. Finally, we reported the pretest risk for psychosis along with the positive (LR+) and negative likelihood ratios (LR-)(for details see\textsuperscript{40}) for the DSM-5-APS assessment.

Since these measures do not tell us whether the DSM-5-APS designation would do more good than harm if used in clinical practice\textsuperscript{41}, we additionally performed net benefit analyses (for conceptual and methodological details see\textsuperscript{41}). Such an approach includes an “exchange rate”, a clinical judgment of the relative value of benefits (such as preventing psychosis in help-seeking individuals) and harms (such as unnecessary treatment) associated with the DSM-5-APS designation. We used 7.69\% as an approximate exchange rate as previously validated\textsuperscript{42} (see eMethods 2). However, as the answers to these kinds of questions and the exchange rates are subjective, we additionally plotted the net benefit for a range of reasonable exchange rates in a decision curve analysis, as recommended\textsuperscript{41}.

*Theoretical impact of DSM-5-APS revisions*
In pilot analyses we tested the theoretical impact of specific revisions of the DSM-5-APS criteria on: diagnostic agreement with the CAARMS 12/2006 (proportion of observed and expected agreement, kappa and P), diagnostic disagreement with the CAARMS 12/2006 (Stuart-Maxwell test\textsuperscript{34, 35}), prognostic measures (2-years Se, Sp, PPV, NPV, computed as indicated above) and prognostic performance of the revised model (Harrell’s C, which is similar to the AUC\textsuperscript{43}).

All analyses were conducted in Stata 13 (STATA Corp., TX, USA) or R version 3.3.0; statistical tests were two-sided and statistical significance was defined as p values of less than 0.05.

RESULTS

Sample characteristics

The sample consisted of 203 consecutive help-seeking individuals accessing the OASIS UHR service during the recruitment period. The majority were low-functioning (average SOFAS=53), Black (43%) and male (65%), with an average age of 25 years (range= 16-37 years) (for details see Table 2).

*** TABLE 2 ABOUT HERE ***

Diagnostic significance: DSM-5-APS vs CAARMS 12/2006

Sociodemographic characteristics associated with DSM-5-APS designation

Upon assessment, 51 individuals were deemed DSM-5-APS+, 91 were DSM-5-APS- and 61 were already psychotic (according to criterion F of the DSM-5-APS). There were no functional, age or sex differences between the three groups. Psychotic
individuals were more likely to be of black ethnicity. There were also significant between-group differences in symptom severity as measured by the CAARMS 12/2006 subscales, with more severe scores in the psychotic group but no significant symptomatic differences between the DSM-5-APS+ and DSM-5-APS- groups.

Diagnostic significance

The percent overall agreement between the DSM-5 (DSM-5-APS+, DSM-5-APS-, Psychosis) and the CAARMS 12/2006 (UHR+, UHR-, Psychosis) was 72.91% (expected agreement by chance 34.33%). The kappa was 0.587 (Z=12.80, p<0.001, 95%CI 0.497-0.677) and the PABAK was 0.594 (95%CI 0.529, 0.658), indicating significant agreement of moderate magnitude. However, the tests for table symmetry ($X^2$=55, P<0.001) and marginal homogeneity ($X^2$=55, P<0.001) were both highly significant, indicating substantial disagreement between the two outcomes. The largest contribution to the symmetry was the fact that 27/105 UHR+ individuals were psychotic on the DSM-5-APS assessment, while there were no psychotic individuals according to the CAARMS 12/2006 who were DSM-5-APS+. Similarly, while there were 27/105 UHR+ individuals who were DSM-5-APS-, there were no DSM-5-APS+ individuals who met the UHR- status. Residual analyses confirmed that half of the UHR+ individuals (54/105) did not receive a DSM-5-APS designation: 23 CAARMS 12/2006 APS (30% of total cases) and 4 GRD (100% of total cases) were classified as DSM-5 APS-, whilst 27 CAARMS 12/2006 BLIPS (100% of total cases) were classified as DSM-5 Psychosis. All BLIPS were not satisfying the DSM-5-APS criterion F: 6 of them were not satisfying the criterion F only and the remaining ones were not satisfying different combinations of criteria beyond criterion F. 3 GRD cases were meeting none of the DSM-5-APS criteria, and 1 was not satisfying criterion E.
Among the 23 APS, 1 was not meeting DSM-5-APS criteria B and D, 4 were not meeting criteria C and E, 3 were not meeting criterion C only, and 15 were not meeting criterion E only (eResults 1).

Agreement between CAARMS 12/2006-defined psychotic cases and DSM-5-APS criterion F was almost perfect. Similarly, there was high agreement between the UHR- and the DSM-5-APS-, except for one UHR- individual who tested psychotic on the DSM-5-APS (in this case, despite not meeting the DSM-5-APS criterion F, the required CAARMS 12/2006 drop in functioning for a BLIPS was not met). Full details are presented in Table 3. Complementary analyses conducted after excluding the GRD and BLIPS cases showed an improved observed diagnostic agreement between DSM-5-APS and CAARMS 12/2006 (86.05%, expected 35.66%), as well as kappa 0.783 (SE=0.054, p<0.001). However, the diagnostic disagreement between the two instruments still persisted (marginal homogeneity $\chi^2=24$, p<0.001).

*** TABLE 3 ABOUT HERE ***

**Prognostic significance: DSM-5-APS vs CAARMS 12/2006**

*Risk of psychosis*

There were 142 non-psychotic cases according to the DSM-5-APS assessment (i.e. all 27 BLIPS and 1 UHR- were excluded): 91 DSM-5-APS- and 51 DSM-5-APS+. The mean follow-up was 532.15 days/1.5 years (SD 275.84 days, range 51-1391 days). There were 17 transitions to psychosis (eResults 2, the clinical fate of BLIPS cases is described elsewhere\textsuperscript{44}), with the last transition observed at 720 days. The median time to transition was 146 days ($25^{th}$-$75^{th}$ percentiles 90-379) in the DSM-5-APS+ group and 106 days ($25^{th}$-$75^{th}$ percentiles 62-453) in the DSM-5-APS- group. The DSM-5-
APS+ group was five times more likely to develop a psychotic disorder than the DSM-5-APS- group: HR 5.379, 95%CI 1.875 -15.430. The Kaplan-Meier failure curve is depicted in Figure 1. The point estimates for the risk of psychosis in the DSM-5-APS+ group were 15.69% (95%CI 8.17% – 28.93%) at 6 months, 17.69% (95%CI 9.62%-31.25%) at 12 months, 20.99% (95%CI 14.80% - 49.50%) at 18 months, and 28.17% (95%CI 14.80%-49.50%) at 21 months (for longer follow-ups there were too few individuals still at risk). The point estimates for the risk of psychosis in the DSM-5-APS- group were 3.30% (95%CI 1.08% – 9.87%) at 6 months, 4.62% (95%CI 1.75% – 11.90%) at 12 months, 6.49% (95%CI 2.68% – 15.27%) at 18 months, and 6.49% (95%CI 2.68% – 15.27%) at 21 and 24 months.

*** FIGURE 1 ABOUT HERE ***

Prognostic accuracy

The pretest risk of psychosis at 24 months was 16.23% (95%CI 9.55 – 28.26). The time dependent AUC of the DSM-5-APS designation in the 142 non-psychotic individuals is depicted in Figure 2, top section. Its point estimates were 64.09 (SE 8.83) at 3 months, 69.95 (SE 7.05) at 6 months, 72.48 (SE 6.66) at 12 months, 72.23 (SE 7.10) at 18 months and 75.69 (SE 6.96) at 24 months. Sensitivity, specificity and PPV/NPV at each timepoint are detailed in eTable 1. The LR+ was 2.262 and the LR- 0.428.

*** FIGURE 2 ABOUT HERE ***

Among these 142 individuals there were no significant differences in the ability to
predict psychosis between the DSM-5-APS and the CAARMS 12/2006 (test of AUCs differences/non-inferiority: \(p>0.05\) at all timepoints), as also shown in the bottom section of Figure 2.

**Net benefit analysis**

At the reference threshold for recommending indicated interventions to prevent psychosis (7.69%), treating on the basis of the DSM-5-APS outcome was associated with a small but significant net benefit of 0.027, compared with treat all. The net benefit increased from this probability threshold (risk of psychosis) onwards, up to 35% risk by two years. The decision curve (Figure 3) confirmed that compared to treating none, treating with indicated interventions on the basis of the DSM-5-APS assessment is associated with net benefits for a range of probability thresholds up to 35% risk by two years. For example, the net benefit of 0.099 at a threshold probability of 15% (Figure 3), indicates that compared to treating none, treating on the basis of the DSM-5-APS assessment is equivalent to a strategy that treats approximately 10 patients at risk for psychosis per hundred patients, without conducting any unnecessary treatment.

*** FIGURE 3 ABOUT HERE ***

**Theoretical impact of DSM-5-APS revisions**

The overall diagnostic agreement between the DSM-5-APS and the CAARMS 12/2006 and the kappa were the highest, and the disagreement the lowest (albeit still significant), when the criterion B was redefined to include cases with psychotic symptoms lasting up to one month and criterion E was conjointly removed (Table 4).
The highest Se and NPV were achieved when all criteria B, C, D, E were removed (but Sp was the lowest). None of the revisions tested showed superior Sp and PPV, compared to the current specifications. Prognostic performance improved when criterion E was removed and was the highest when the criterion C was removed.

*** TABLE 4 ABOUT HERE ***

DISCUSSION

In 203 consecutive help-seeking individuals accessing UHR services on suspicion of psychosis risk, the agreement between the DSM-5-APS and the CAARMS 12/2006 was significant but only moderate in magnitude (kappa 0.59). Disagreement mostly involved UHR+ individuals under the CAARMS 12/2006, as half of them (one third of APS cases and all of BLIPS or GRD cases) retained a different designation under the DSM-5-APS. Across 142 non-psychotic cases followed up for a mean of 1.5 years, those meeting the DSM-5-APS criteria had a fivefold probability (HR 5.379) of developing a psychotic disorder compared to those not meeting these criteria, with a 21-month cumulative risk of psychosis of 28% and 6%, respectively. The AUC of the DSM-5-APS was acceptable (0.76 at 24 months), with high NPV and modest PPV, and comparable to the AUC of the CAARMS 12/2006 at any timepoint. Decision curve analysis suggested that in non-psychotic (and non-BLIPS) individuals, the DSM-5-APS designation may be clinically useful within a 7%-35% range of threshold probability (at two-years). Removal of criterion E and C of the DSM-5-APS may improve its prognostic performance, transdiagnostic value and clinical implementation.
This is the first study to address the diagnostic significance of the DSM-5-APS in clinical services for UHR individuals, using the CAARMS 12/2006 as gold standard reference. Our prospective cohort design that included all consecutive individuals seeking help at UHR services and undergoing an assessment on suspicion of psychosis risk allowed estimation of the real-world diagnostic and prognostic significance of the DSM-5-APS in these populations. The sociodemographic characteristics of the current sample and its substantial pretest psychosis risk enrichment (16% at two years) are key characteristics of selected populations seeking help at UHR services worldwide. In this clinical context, we found a significant diagnostic agreement (higher than that observed by chance) between the DSM-5-APS and the CAARMS 12/2006, although only moderate in magnitude (kappa 0.59). Agreement was optimal only for psychotic cases identified by the CAARMS 12/2006, suggesting that the DSM-5-APS has a lower psychotic threshold than the CAARMS 12/2006. Significant disagreement was indeed present and mostly due to intake criteria differences, as around half of the UHR+ individuals according to the CAARMS 12/2006 were not at risk as per the DSM-5-APS. Not surprisingly, all BLIPS cases and the few GRD cases were not meeting DSM-5-APS criteria. More importantly, around one third of individuals meeting the APS subgroup on the CAARMS 12/2006 did not actually meet DSM-5-APS criteria, mostly because they were presenting with comorbidities that were not satisfying criterion E of the DSM-5-APS. Conversely, because of the functional decline required by the CAARMS 12/2006 (for a meta-analysis on the impact of functioning in UHR samples see ), most APS individuals satisfied criterion D of the DSM-5-APS. The substantial disagreement between the attenuated psychosis symptoms defined psychometrically (CAARMS 12/2006) and clinically (DSM-5-APS) persisted even when GRD and
BLIPS were excluded. This indicates that the two instruments are not diagnostically interchangeable and that the use of CAARMS 12/2006 with the aim of defining DSM-5-APS cases\textsuperscript{17, 18} is not psychometrically sound (but see below re: prognostic non-inferiority). These discrepancies may be of clinical relevance, as the DSM-5-APS is considered a disorder per se, and therefore it is a broad marker for identifying individuals who merit active care to address a variety of symptoms, independent from their longitudinal risk of developing psychosis\textsuperscript{47}.

This is also the first prospective study to test the actual prognostic significance of the DSM-5-APS designation for psychosis onset in the pool of individuals who were not DSM-5-APS psychotic at baseline, excluding also all BLIPS individuals. BLIPS cases are problematic as they have a distinctive diagnostic and prognostic profile as compared to the APS and GRD subgroups, as extensively demonstrated by our team\textsuperscript{4, 30, 44, 48, 49}. We have clearly shown that, over an average follow-up of 1.5 years, DSM-5-APS+ individuals were five times more at risk of developing psychosis (HR 5.379), than the DSM-5-APS- individuals. These findings indicate that DSM-5-APS, if used in non-psychotic patients accessing UHR services, could effectively predict conversion to psychosis. Its ability to rule in psychosis appeared slightly better than that reported for the UHR psychometric instruments. The meta-analytical risk of psychosis of the CAARMS/SIPS-defined APS group (24-month mean 19%, 95%CI 15% - 23%)\textsuperscript{30} and the LR+ of the UHR instruments (1.82 at 38-months)\textsuperscript{6} seem lower than those reported here for the DSM-5-APS (21-month average risk of psychosis 28.17% and LR+ 2.262). Conversely, the DSM-5-APS appeared slightly less accurate in predicting the absence of psychosis than the psychometric UHR assessment. The meta-analytical risk of psychosis of the CAARMS/SIPS-defined APS group (24-
month mean 1%, 95% CI 0%-3%)\(^3\) and the LR- of the UHR instruments (0.09 at 38-months)\(^6\) seem lower (thus better for excluding psychosis), compared to the DSM-5-APS (21-month average risk of psychosis 6.49%, LR- 0.428). As these differences were counterbalanced, the overall prognostic accuracy of the DSM-5-APS was non-inferior and almost comparable with the CAARMS 12/2006 at any timepoint. On a conceptual side, this finding may appear surprising, given the only moderate diagnostic agreement (kappa) between the two instruments. Of interest, prognostic accuracy is also comparable across the CAARMS and the SIPS\(^6\), despite their substantial diagnostic differences\(^28\). These observations corroborate the notion that the actual prognostic accuracy of all of these instruments is mostly determined by their pre-test risk enrichment, as our group has extensively shown\(^8, 9, 40, 45, 50\). Since we found a 16% 2-year pretest risk of psychosis in this help-seeking sample, it is likely that in general non-help seeking populations the DSM-5-APS is associated with substantially lower transition risks. Consequently, most of the findings reported here are not generalizable outside help-seeking samples accessing UHR services.

On a pragmatic side, the finding of non inferiority of DSM-5-APS for the prediction of psychosis in UHR services, compared to standard psychometric interviews, may have some practical clinical relevance, although further replication studies are required. For example, we qualitatively estimated that a DSM-5-APS assessment is relatively quicker to administer in UHR services, requiring on average 23 minutes, compared with 120 minutes for the CAARMS 12/2006\(^26\) (but see the elimitations) and recent studies show that the label DSM-5-APS is equally accepted compared to the standard “UHR” or “At Risk Mental State”\(^51\).

Our study was also the first to directly investigate the potential clinical usefulness of
the DSM-5-APS designation for clinical practice in UHR services. We used net benefit and decision curve analysis approaches -that were originally developed in cancer research- to investigate the clinical usefulness of different strategies for offering indicated interventions to UHR patients, taking into consideration both harms and benefits. We found that in non-BLIPS cases, indicated interventions selected on the basis of the DSM-5-APS outcomes were associated with significant net benefits for a range of psychosis risk (at 2-years), spanning from 7% to 35%. Such a range is clinically meaningful for UHR services since it is unlikely an UHR individual would need a clinical assessment to receive indicated interventions when his/her 2-year risk of psychosis is lower or greater than this range. These findings may be of particular relevance for the continued evaluation of the DSM-5-APS’s clinical validity and for its appropriateness for inclusion in the next update of DSM-5.112.

We also tested in pilot analyses the first theoretical diagnostic and prognostic impact of some potential revisions of the DSM-5-APS. None of them had a dramatic effect in improving the prognostic performance, presumably in light of the fact that the observed risk in these samples is mostly accounted at pre-test level (i.e. 16% risk at 2-years), through the recruitment strategies adopted. However, our analyses suggested that removing either criteria E and C may be associated with an improved prognostic performance. The removal of criterion E, which poses substantial differential diagnostic challenges to clinicians, would additionally facilitate the transdiagnostic potential of the DSM-5-APS and better fill the current implementation gap for the identification of at risk individuals (only 5% of at risk individuals in secondary mental health care are currently detected by UHR services)42. For example, since the UHR criteria have been successfully applied to individuals affected with 22q11 deletion
syndrome\textsuperscript{52} or may be used in high potency cannabis users\textsuperscript{6} a revision of the DSM-5- APS should allow inclusion of these samples\textsuperscript{53}. However, because of the limited sample size, these analyses should be considered cautiously. Future studies may leverage on our methodological approach and validate these findings. Limitations of the study are outlined in the eLimitations.

CONCLUSIONS

In help-seeking individuals accessing UHR services, the DSM-5-APS and the CAARMS 12/2006 definitions of attenuated psychosis symptoms are diagnostically different, but prognostically comparable. The DSM-5-APS designation offers some promising clinical usefulness for the prevention of psychosis that may be improved by future revisions.
REFERENCES:


## Table 1. Competing diagnostic constructs in services for individuals at ultra high risk for psychosis: DSM-5-APS vs CAARMS Version 12/2006

<table>
<thead>
<tr>
<th></th>
<th>DSM-5-APS&lt;sup&gt;34&lt;/sup&gt;</th>
<th>CAARMS Version 12/2006&lt;sup&gt;35&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symptoms</strong></td>
<td>At least one of (criterion A): 1. Delusions 2. Hallucinations 3. Disorganized speech</td>
<td>P1 - Unusual Thought Content  P2 - Non Bizarre Ideas  P3 - Perceptual Abnormalities  P4 - Disorganized Speech</td>
</tr>
</tbody>
</table>
| **Severity and frequency** | “Symptoms present in attenuated form, with relatively intact reality testing, and are of sufficient severity and/or frequency to warrant clinical attention (criterion A)” | APS<sup>3</sup>:  
*Subthreshold intensity*
Severity score of 3–5 on P1, 3–5 on P2, 3–4 on P3 and/or 4–5 on P4 PLUS Frequency score* of 3–6 on P1, P2, P3 and/or P4  
*Subthreshold frequency*
Severity score of 6 on at least one of P1, P2, P4 and/or 5–6 on P3 PLUS Frequency score of 3 on P1, P2, P3 and/or P4  
GRD<sup>2</sup>:  Schizotypal Personality Disorder in identified patient OR Family history of psychosis in first degree relative  
BLIPS<sup>3</sup>:  Severity score of 6 on at least one of P1, P2, P4 and/or 5–6 on P3 PLUS Frequency score of 4–6 on P1, P2, P3 and/or P4<sup>55</sup> PLUS Lasting up to 7 days |
| **Recency**    | “Symptoms must be present at least once per week for the past month (criterion B)” | The highest level of symptom is rated, even if it is not present at interview or in the past month |
| **Onset**      | “Symptoms must have begun or worsened in the past year (criterion C)” | Symptoms should have been present in the previous 12 months and for not longer than 5 years |
| **Level of functioning** | “Symptoms are sufficiently distressing and disabling to the individual to warrant clinical attention (criterion D)” | 30% drop in SOFAS score from premorbid level, sustained for a month, within past 12 months OR SOFAS score<50 for past 12 months or more |
| **Exclusion criteria** | “Symptoms are better explained by another mental disorder, including a depressive or bipolar disorder with psychotic features (criterion E)” | Symptoms occur only during peak intoxication from a substance known to be associated with psychotic experiences (e.g. hallucinogens, amphetamines, cocaine) |
| **Psychosis threshold** | “Clinical criteria for a psychotic disorder have never been met (criterion F)” | Severity score of 6 on at least one of P1, P2, P4 and/or 5–6 on P3 PLUS Frequency score of 4–6 on P1, P2, P3 and/or P4 |
| **Transition to psychosis** | Unknown | 27.4% [95% CI, 24.6%-30.4%] at 31 months<sup>36</sup> |

*0 – Absent; 1 – Less than 1/mo; 2 – 1/mo to 2/wk - < 1 h per occasion; 3 – 1/mo to 2/wk > 1 h per occasion OR 3 to 6/wk - < 1 h per occasion; 4 – 3 to 6/wk - > 1 h per occasion OR Daily - < 1 h per occasion; 5 – Daily - > 1 h per occasion OR several times/d; 6 – Continuous.

<sup>1</sup> APS, Attenuated Psychosis Group; <sup>2</sup> GRD, Vulnerability Group; <sup>3</sup> BLIPS, Brief Limited Intermittent Psychotic Symptoms Group.
Table 2. Sociodemographic and clinical characteristics of individuals undergoing the assessment at the OASIS service

<table>
<thead>
<tr>
<th></th>
<th>Total sample (n=203)</th>
<th>DSM-5-APS+ (n=51)</th>
<th>DSM-5-APS- (n=91)</th>
<th>DSM-5-APS Psychosis (n=61)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Age (years)</td>
<td>25.14</td>
<td>5.39</td>
<td>24.84</td>
<td>5.54</td>
</tr>
<tr>
<td>Functional level (SOFAS)</td>
<td>52.72</td>
<td>13.79</td>
<td>52.89</td>
<td>10.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
<th>X2</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>132</td>
<td>65.02</td>
<td>34</td>
<td>66.67</td>
<td>57</td>
<td>62.64</td>
<td>41</td>
<td>67.21</td>
<td>0.417</td>
<td>0.816</td>
</tr>
<tr>
<td>Females</td>
<td>71</td>
<td>34.98</td>
<td>17</td>
<td>33.33</td>
<td>34</td>
<td>37.36</td>
<td>20</td>
<td>32.79</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Self assigned ethnicity   |       |       |       |       |       |       |       |       |     |     |
| Black                     | 87    | 42.86 | 22    | 43.14 | 29    | 31.87 | 36    | 59.02 | 26.03 | <0.001 |
| White                     | 84    | 41.38 | 16    | 31.37 | 50    | 54.95 | 18    | 29.51 |      |     |
| Asian                     | 10    | 4.93  | 4     | 7.84  | 3     | 3.30  | 3     | 4.92  |      |     |
| Mixed                     | 13    | 6.40  | 8     | 15.69 | 3     | 3.30  | 2     | 3.28  |      |     |
| Other                     | 9     | 4.43  | 1     | 1.96  | 6     | 6.59  | 2     | 3.28  |      |     |

| CAARMS 12/2006            |       |       |       |       |       |       |       |       |     |     |
| Severity of unusual thought content | 2 | 1-5 | 3 | 2-4 | 3 | 2-4 | 5 | 4-6 | 64.86 | <0.001 |
| Severity of non bizarre ideas | 3 | 1-5 | 3 | 1-5 | 3 | 2-5 | 6 | 4.5-6 | 72.97 | <0.001 |
| Severity of perceptual abnormalities | 3 | 0-5 | 3 | 1-5 | 3 | 1-5 | 5 | 2.5-6 | 29.98 | <0.001 |
| Severity of disorganized speech | 2 | 0-3 | 2 | 0-3 | 2 | 0-3 | 3 | 0.5-4 | 17.59 | <0.001 |

SD, standard deviation; X2 Fisher exact test; IQR, 25°-75° interquartile range, K-W Kruskal-Wallis test.
Table 3. Diagnostic comparison between DSM-5-APS and CAARMS 12/2006 in individuals seeking help at the OASIS service.

<table>
<thead>
<tr>
<th>CAARMS 12/2006 Outcome</th>
<th>DSM-5-APS Outcome</th>
<th>APS+</th>
<th>APS-</th>
<th>Psychosis</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>UHR+ APS</td>
<td>Count</td>
<td>51</td>
<td>23</td>
<td>0</td>
<td>74</td>
</tr>
<tr>
<td>%</td>
<td>68.92</td>
<td>31.08</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>18.6</td>
<td>33.2</td>
<td>22.2</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Adjusted Residuals</td>
<td>10.89</td>
<td>-2.98</td>
<td>-7.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIPS</td>
<td>Count</td>
<td>0</td>
<td>0</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>%</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>6.8</td>
<td>12.1</td>
<td>8.1</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Adjusted Residuals</td>
<td>-3.23</td>
<td>-5.03</td>
<td>8.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRD</td>
<td>Count</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>%</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>1</td>
<td>1.8</td>
<td>1.2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Adjusted Residuals</td>
<td>-1.17</td>
<td>2.24</td>
<td>-1.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UHR-</td>
<td>Count</td>
<td>0</td>
<td>64</td>
<td>1</td>
<td>65</td>
</tr>
<tr>
<td>%</td>
<td>0</td>
<td>98.46</td>
<td>1.54</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>16.3</td>
<td>29.14</td>
<td>19.53</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Adjusted Residuals</td>
<td>-5.66</td>
<td>10.55</td>
<td>19.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychosis</td>
<td>Count</td>
<td>0</td>
<td>0</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>%</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>8.3</td>
<td>14.79</td>
<td>9.92</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Adjusted Residuals</td>
<td>-3.64</td>
<td>-5.66</td>
<td>9.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>Count</td>
<td>51</td>
<td>91</td>
<td>61</td>
<td>203</td>
</tr>
<tr>
<td>%</td>
<td>25.12</td>
<td>44.83</td>
<td>30.05</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

CAARMS, Comprehensive Assessment of At Risk Mental States. Adjusted residuals smaller than -3.10 or greater than 3.10 indicated a statistically significant difference at P<0.0001 of observed counts versus counts expected by chance. However, three GRD cells have expected count less than 5 and are likely to be underpowered. Fisher’s exact test P<0.001.
Figure 1. Kaplan-Meier failure function for the risk of developing psychosis in non-psychotic individuals meeting (n=51) and not meeting (n=91) the DSM-5-APS criteria. There were 4 cumulative failures during the first 3 months, 8 during the first 6 months, 12 in the first 12 months, 15 in the first 18 months, 17 in the first 24 months.
Figure 2. Prognostic accuracy measure (Area Under the Curve, AUC) for the DSM-5-APS designation over time in 142 non psychotic help-seeking individuals undergoing assessment at UHR services. Top part of the figure: AUC (red line) with 95% CI confidence bands (dotted red lines) over follow-up time, the 0.5 line indicates non-significant results. Bottom part of the figure: AUC difference between DSM-5-APS vs CAARMS 12/2006 (AUCΔ, blue line), with 95% CI confidence bands (dotted blue lines) over follow-up time; the 0 line indicates non-significant differences.
Figure 3. Decision curve analysis showing the clinical usefulness (net benefits) of primary indicated interventions implemented on the basis of the DSM-5-APS designation (green line), compared to treating all non-psychotic (and non BLIPS) patients accessing UHR services (blue line) or to treating no patients at all (red line) for a range of threshold probability (risk of psychosis at 2-year).
Table 4. Theoretical impact of criteria revisions on the diagnostic and prognostic significance of DSM-5-APS in individuals undergoing assessment for psychosis risk at UHR services

<table>
<thead>
<tr>
<th>Suggested revisions of DSM-5-APS criteria</th>
<th>N</th>
<th>Diagnostic agreement with CAARMS 12/2006</th>
<th>Diagnostic disagreement with CAARMS 12/2006</th>
<th>N</th>
<th>Se (2yr)</th>
<th>Sp (2yr)</th>
<th>PPV (2yr)</th>
<th>NPV (2yr)</th>
<th>Prognostic performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Agreement / Expected</td>
<td>Kappa</td>
<td>P</td>
<td>X²</td>
<td>P</td>
<td>% (SE)</td>
<td>% (SE)</td>
<td>% (SE)</td>
</tr>
<tr>
<td><strong>Current DSM-5-specifications</strong></td>
<td>203</td>
<td>72.91% / 34.33%</td>
<td>0.587</td>
<td>&lt;0.001</td>
<td>55</td>
<td>&lt;0.001</td>
<td>142</td>
<td>72.80 (11.54)</td>
<td>78.57 (7.78)</td>
</tr>
<tr>
<td>Removal of criterion B: “Symptoms must be present at least once per week for the past month”</td>
<td>203</td>
<td>72.91% / 34.33%</td>
<td>0.587</td>
<td>&lt;0.001</td>
<td>55</td>
<td>&lt;0.001</td>
<td>142</td>
<td>72.80 (11.54)</td>
<td>78.57 (7.78)</td>
</tr>
<tr>
<td>Removal of criterion C: “Symptoms must have begun or worsened in the past year”</td>
<td>203</td>
<td>73.89% / 32.62%</td>
<td>0.612</td>
<td>&lt;0.001</td>
<td>48</td>
<td>&lt;0.001</td>
<td>142</td>
<td>72.80 (11.54)</td>
<td>78.57 (7.78)</td>
</tr>
<tr>
<td>Removal of criterion D: “Symptoms are sufficiently distressing and disabling to the individual to warrant clinical attention”</td>
<td>203</td>
<td>72.91% / 34.33%</td>
<td>0.587</td>
<td>&lt;0.001</td>
<td>55</td>
<td>&lt;0.001</td>
<td>142</td>
<td>72.80 (11.54)</td>
<td>78.57 (7.78)</td>
</tr>
<tr>
<td>Removal of exclusion criterion E: “Symptoms are better explained by another mental disorder” and are “attributable to the physiological effects of a substance or another medical condition”</td>
<td>203</td>
<td>77.83% / 34.37%</td>
<td>0.662</td>
<td>&lt;0.001</td>
<td>29</td>
<td>&lt;0.001</td>
<td>142</td>
<td>82.56 (9.96)</td>
<td>53.57 (9.46)</td>
</tr>
<tr>
<td>Removal of criterion E and C</td>
<td>203</td>
<td>77.34% / 35.82%</td>
<td>0.647</td>
<td>&lt;0.001</td>
<td>33</td>
<td>&lt;0.001</td>
<td>142</td>
<td>86.9 (9.2)</td>
<td>50 (9.48)</td>
</tr>
<tr>
<td>Removal of criteria B, C, D, E</td>
<td>203</td>
<td>75.25% / 36.28%</td>
<td>0.612</td>
<td>&lt;0.001</td>
<td>39</td>
<td>&lt;0.001</td>
<td>142</td>
<td>90.84 (8.71)</td>
<td>50 (9.48)</td>
</tr>
<tr>
<td>Redefinition of criterion B: “Clinical criteria for a psychotic disorder lasting more than 1 month have never been met”</td>
<td>203</td>
<td>75.86% / 34.91%</td>
<td>0.629</td>
<td>&lt;0.001</td>
<td>49</td>
<td>&lt;0.001</td>
<td>169</td>
<td>64.21 (11.03)</td>
<td>76.19 (6.59)</td>
</tr>
<tr>
<td>Redefinition of criterion B (as above) and removal of criterion E</td>
<td>203</td>
<td>82.76% / 37.44%</td>
<td>0.724</td>
<td>&lt;0.001</td>
<td>15</td>
<td>0.002</td>
<td>169</td>
<td>76.57 (9.72)</td>
<td>54.76 (7.70)</td>
</tr>
</tbody>
</table>