Appendix 2: Levels of evidence and grades of recommendation used to describe the strength of recommendations in clinical practice guidelines (CPG) addressing the pharmacological treatment of first episode schizophrenia.

PORT 2009	Spain 2009	Malaysia, 2009	Singapore 2011	BAP 2011	WFSBP, 2012	SIGN, 2013	Harvard 2013	NICE 2014	RANZCP, 2016
Must have at least 2 RCTs to make a recommendation	Ia Meta-analysis of RCTs Ib At least one RCT IIa At least one well designed non-randomised controlled prospective study IIb At least one well designed quasi-experimental study III Well designed observational studies eg comparative study, correlation study or case-control studies IV Expert opinion and clinical experience Grade A: Evidence level 1a or 1b. At least one good quality RCT. Grade B: Evidence level IIa, IIb, or III. Methodologically correct clinical trials that are not RCTs Grade C: Evidence level IV. Expert opinion in the absence of other clinical evidence.	Level 1, good strength, Meta-analysis of RCT, systematic review. Level 2, good strength. Large sample RCT Level 3, Good to fair strength. Small sample RCT. Level 4, Good to fair strength. Non-randomised controlled prospective trial. Level 5, fair strength. Non-randomised controlled prospective trial with historical control. Level 6. Fair strength. Cohort study. Level 7, Poor strength, case-controlled study. Level 8, Poor strength, Non-controlled clinical series, descriptive studies multi-centre Level 9, poor strength, Expert committees, consensus, case reports, anecdotes.	1++ High quality meta- analysis, systematic reviews of RCTs or RCT with very low risk of bias. 1+ Well-conducted meta-analysis, systematic reviews of RCTs or RCTS with a low risk of bias 1- Meta-analysis, systematic reviews of RCTs or RCTs with a high risk of bias 2++ High quality systematic reviews of case control or cohort studies, High quality case control or cohort studies with a very low risk of bias or confounding and a high probability that the relationship is causal 2+ Well conducted case control or cohort studies with a low risk of bias or confounding and a moderate probability that the relationship is causal. 2- Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal.	***Causal relationships and treatment*** Category I; Meta-analysis of RCTs, at least one large good quality RCT or replicated, smaller RCTs. Category II: Small non-replicated RCT; at least one controlled study or at least one other quasi experimental study. RCT must have a control treatment arm. Category III: non-experimental descriptive studies eg comparative, correlation or case control. Category (IV) Expert committee report/ opinion/ clinical experience ***Non-causal relationships*** Category I: Evidence from large representative population samples. Category II: Evidence from small, well-designed, but not necessarily representative samples.	Category of Evidence: A: Full evidence from controlled studies: Two or more double blind RCT vs placebo and one or more RCT vs active comparator with placebo arm or well conducted non-inferiority trial. If there is an existing negative study it must be outweighed by at least 2 positive studies or a meta-analysis. B: Limited positive evidence from controlled studies. One or more RCT showing superiority to placebo or RCT vs comparator without placebo control and no negative studies exist. C Evidence from Uncontrolled studies/ case reports/ expert opinion. C1: Uncontrolled studies: 1 or more positive naturalistic study, comparison with an existing drug with sufficient sample size and no negative studies. C2: Case reports. One or more positive case	1++ High quality meta- analysis, systematic reviews of RCTs or RCT with very low risk of bias. 1+ Well-conducted meta-analysis, systematic reviews of RCTs or RCTS with a low risk of bias 1- Meta-analysis, systematic reviews of RCTs or RCTs with a high risk of bias 2++ High quality systematic reviews of case control or cohort studies, High quality case control or cohort studies with a very low risk of bias or confounding and a high probability that the relationship is causal 2+ Well conducted case control or cohort studies with a low risk of bias or confounding and a moderate probability that the relationship is causal. 2- Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal.	None described	Strength of recommendation described in the language of the recommendation. <u>Must or must not</u>: Legal duty to apply recommendation of if consequences of not following recommendation are serious or life threatening. <u>Should or should not:</u> Indicates a strong recommendation. 'Offer', 'refer', 'advise' when confident that for the vast majority of patients an intervention will do more good than harm and be cost effective. Conversely 'do not offer' when confident that intervention will not be of benefit for most patients. <u>Could be used.</u> 'Consider' if confident that an intervention will do more good than harm for most patients, be cost effective but other options may be similarly cost effective. Choice of the intervention more likely to	Recommendations are either Evidence based (EBR) or consensus based (CBR). The level of evidence on which EBR is according to the National Health and Medical Research Council's levels of evidence for healthcare interventions. Level I: A systematic review of level II studies. Level II: A randomised controlled trial. Level III-1: A pseudo-randomised controlled trial. Level III-2: A comparative study with concurrent controls: non-randomised, experimental trial. Cohort studies. Case-control study. Interrupted time-series with a control group. Level III-3: A comparative study without concurrent controls. Historical control study. Two or more single-arm studies. Interrupted time series without

		Grades of Recommendation. A. At least one meta-analysis, systematic review, RCT, or evidence rated as good and directly applicable to the target population. B. Evidence from well conducted clinical trials, directly applicable to the target population, and demonstrating overall consistency of results; or evidence extrapolated from meta-analysis, systematic review, or RCT. C. Evidence from expert committee reports, or opinions and/or clinical experiences of related authorities; indicates absence of directly applicable clinical studies of good quality.	3 Non-analytic studies eg case reports, case series 4 Expert opinion **Grades of Recommendation.** **A** At least one meta-analysis, systematic review of RCTs, or RCT rated as 1++ and directly applicable to the target population; or a body of evidence consisting principally of studies rated as 1+ applicable to target population and demonstrating overall consistency of results. **B** A body of evidence consisting principally of studies rated as 2++ applicable to target population and demonstrating overall consistency of results; or extrapolated evidence from studies rated as 1++ or 1+ **C** A body of evidence consisting principally of studies rated as 2+ applicable to target population and demonstrating overall consistency of results; or extrapolated evidence from studies rated as 2+	Category III: Evidence from non-representative surveys, case reports. Category IV: Evidence from expert committee reports or opinions and /or clinical opinions of respected authorities. ***Strength of recommendation*** A: Category I B Category II or extrapolated from category I C: Category III or extrapolated from category I or II D: Category IV or extrapolated from category I, II or III S: Standard of good practice	reports. No negative controlled studies. C3: Expert opinion or clinical experience. D: Inconsistent results. Equal number of positive and negative RCTs E Negative evidence. Majority of RCTs show no benefit over placebo or comparator medication. F: Lack of Evidence. **Grades of recommendation:** 1: Category A plus good risk benefit ratio. 2: Category A and moderate risk-benefit ratio 3: Category B 4: Category C 5: Category D	3 Non-analytic studies eg case reports, case series 4 Expert opinion **Grades of Recommendation.** **A** At least one meta-analysis, systematic review of RCTs, or RCT rated as 1++ and directly applicable to the target population; or a body of evidence consisting principally of studies rated as 1+ applicable to target population and demonstrating overall consistency of results. **B** A body of evidence consisting principally of studies rated as 2++ applicable to target population and demonstrating overall consistency of results; or extrapolated evidence from studies rated as 1++ or 1+ **C** A body of evidence consisting principally of studies rated as 2+ applicable to target population and demonstrating overall consistency of results; or extrapolated evidence from studies rated as 2+		depend on the patient values and preferences and so more consultation should take place. System above does not apply to 2009 recommendations.	a parallel control group. Level IV: Case series with either post-test or pre-test/ post-test outcomes.

			D Evidence level 3 or 4 or extrapolated evidence from studies rated as 2+ **GPP** (Good Practice Point) Recommended best practice based on clinical experience of guideline development group.			**D** Evidence level 3 or 4 or extrapolated evidence from studies rated as 2+ **GPP** (Good Practice Point) Recommended best practice based on clinical experience of guideline development group.			