

Supplement Table 2. Risk of bias summary for included study according to ROBINS-1

Author (year)	Bias due to confounding	Bias in selection of participants into study	Bias in classification of interventions	Bias due to deviations from intended intervention	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result
Büchter et al (1998) [15]	Serious The initial testicular volumes were not well adjusted	NI No information about the patient inclusion	Low Both therapies were well-defined intervention	Moderate One patient received GnRH therapy right after gonadotropin therapy	Low No losses were reported	Moderate No masking strategy specified	Low All results are subjective and well-defined measure
Gong et al (2015) [10]	Low All baseline characteristics were balanced between two groups	NI No information about the patient inclusion	Low Both therapies were well-defined intervention	Low NO co-intervention between intervention groups	Low No losses were reported	Moderate No masking strategy specified	Low All results are subjective and well-defined measure
Huang et al (2015) [13]	Low All baseline characteristics were balanced between two groups	Low Consecutive patients were included	Low Both therapies were well-defined intervention	Low NO co-intervention between intervention groups	Low No losses to FU	Moderate No masking strategy specified	Low All results are subjective and well-defined measure
Kliesch et al (1994) [16]	Low All baseline characteristics were balanced between two groups	NI No information about the patient inclusion	Low Both therapies were well-defined intervention	Moderate One patient received GnRH therapy right after gonadotropin therapy	Low No losses to FU	Moderate No masking strategy specified	Low All results are subjective and well-defined measure
Liu et al (1988) [11]	Serious Not adjusted for baseline variables (e.g., age)	NI No information about the patient inclusion	Low Both therapies were well-defined intervention	Low NO co-intervention between intervention groups	Serious Samples of two patients were unable to be collected	Moderate No masking strategy specified	Low All results are subjective and well-defined measure
Mao et al (2017) [14]	Serious Not adjusted for baseline variables (e.g., age, testicular volume)	Serious Only subject with more than 12 month FU were recruited	Low Both therapies were well-defined intervention	Low NO co-intervention between intervention groups	Low No losses to FU	Moderate No masking strategy specified	Low All results are subjective and well-defined measure
Schopohl et al (1991) [12]	Low All baseline characteristics were balanced between two groups	NI No information about the patient inclusion	Low Both therapies were well-defined intervention	Low One patient changed from GnRH to gonadotropin therapy, but did not get a spermatogenesis and the bias is minimal	Low No losses to FU	Moderate No masking strategy specified	Low All results are subjective and well-defined measure

NI: no information, GnRH: gonadotropin-releasing hormone, FU: follow-up.