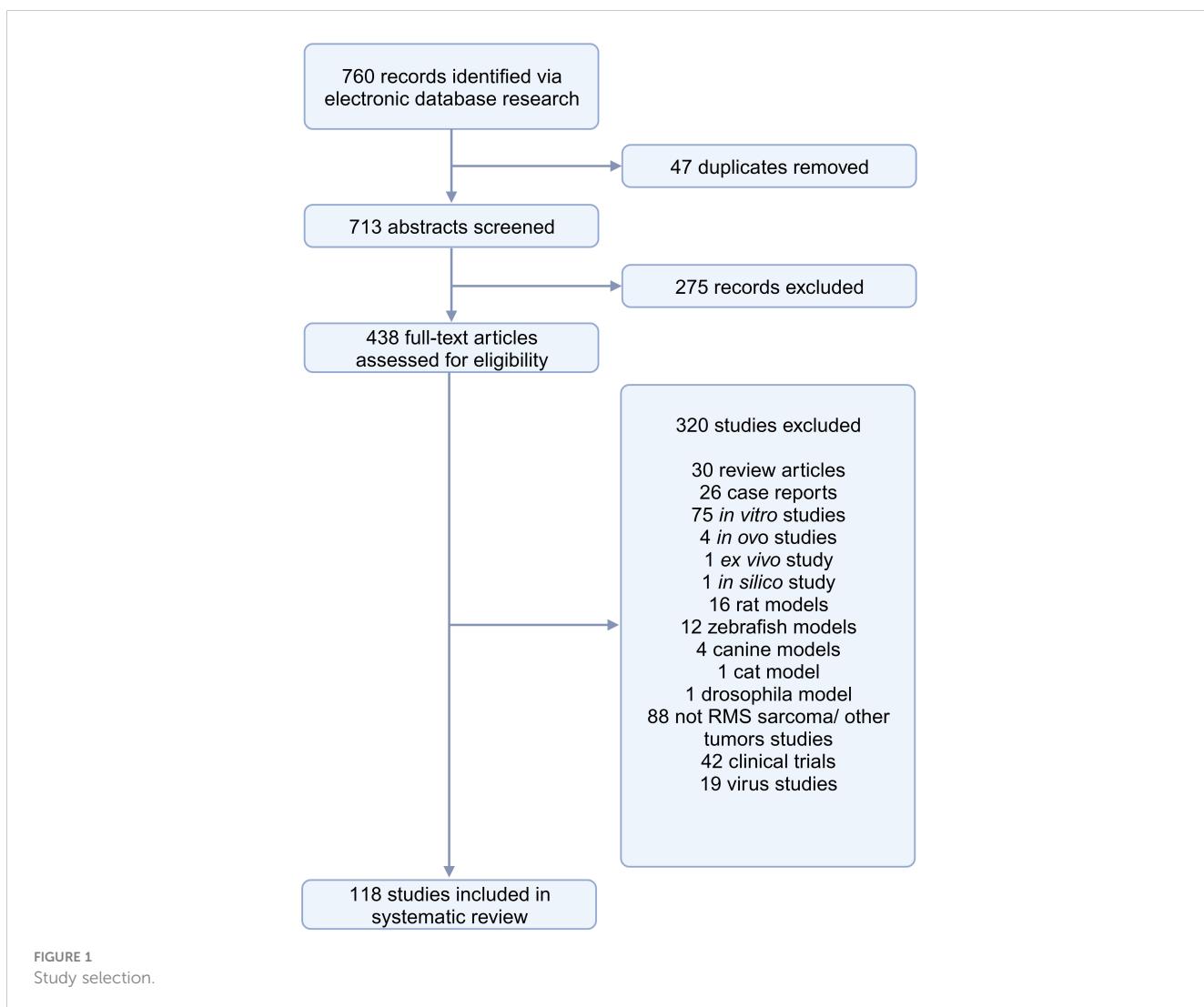


e.g., new drug testing on ectopic CDX), moderate (distant translation possible, e.g., establishment of PDX tumor bank), high (direct translational impact, e.g., MRI/histology-correlation studies, studies on long-term effects after radiotherapy), very high (immediate translational impact by using complex mouse systems, e.g., Single Mouse Trials). All descriptive statistics were done using Jamovi (Version 2.3.16) and GraphPad Prism (Version 10.0.3) software. The choropleth map was created using Datawrapper (<https://www.datawrapper.de>).

Results

The search conducted in PubMed and Web of Science yielded a total of 713 unique papers after removing 47 duplicates. After title and abstract screening, 275 papers were excluded, leaving 438 articles (Figure 1). Among these, 118 articles (26.9%) met the eligibility criteria and were considered for qualitative synthesis. Supplementary Table 1 contains the comprehensive list of included studies. Supplementary Figure 1 illustrates the geographical coverage of the data, encompassing 19 countries across 4

continents. The results of the risk of bias within the included studies are reported in Figure 2. Besides providing information on ethical statements, the majority of studies offered insufficient information regarding selection bias (sequence generation, baseline characteristics, and allocation concealment) and performance bias (random housing and investigator blinding). The mean sample size of animals per study was 65.1 ± 107.8 , with a range from 3 to 499 (data available for $n = 21$ studies, 17.8%), and the mean sample size per treatment group was 7.9 ± 2.6 , with a range from 4 to 12 (data available for $n = 26$ studies, 22.1%). Among different RMS mouse systems utilized in our study sample ($n = 118$), cell line-derived xenografts (CDX) emerged as the most commonly used ($n = 75$, 63.6%), followed by patient-derived xenografts (PDX) and syngeneic models, each accounting for 11.9% ($n = 14$). A smaller subset of studies utilized exclusively genetically engineered mouse models (GEMM) ($n = 7$, 5.9%). There were also combinations of different model categories, such as PDX/CDX or GEMM/syngeneic, accounting for 5.9% ($n = 7$). Additionally, one study employed a virus-induced RMS model, representing 0.8% of the total studies. The summary of various existing mouse model systems utilized in RMS research is presented in Figure 3.



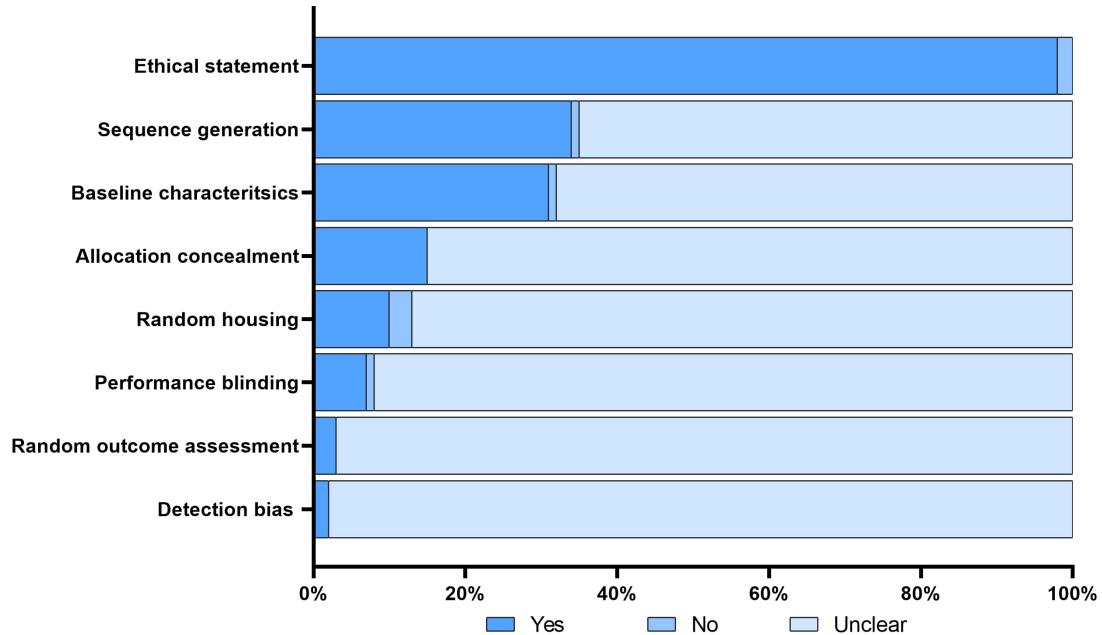


FIGURE 2

Risk of bias. The figure illustrates the risk of bias assessment for each included study using the modified SYRACLE tool. A 'yes' score indicates a low risk of bias; a 'no' score indicates a high risk of bias; and an 'unclear' score indicates an unknown risk of bias.

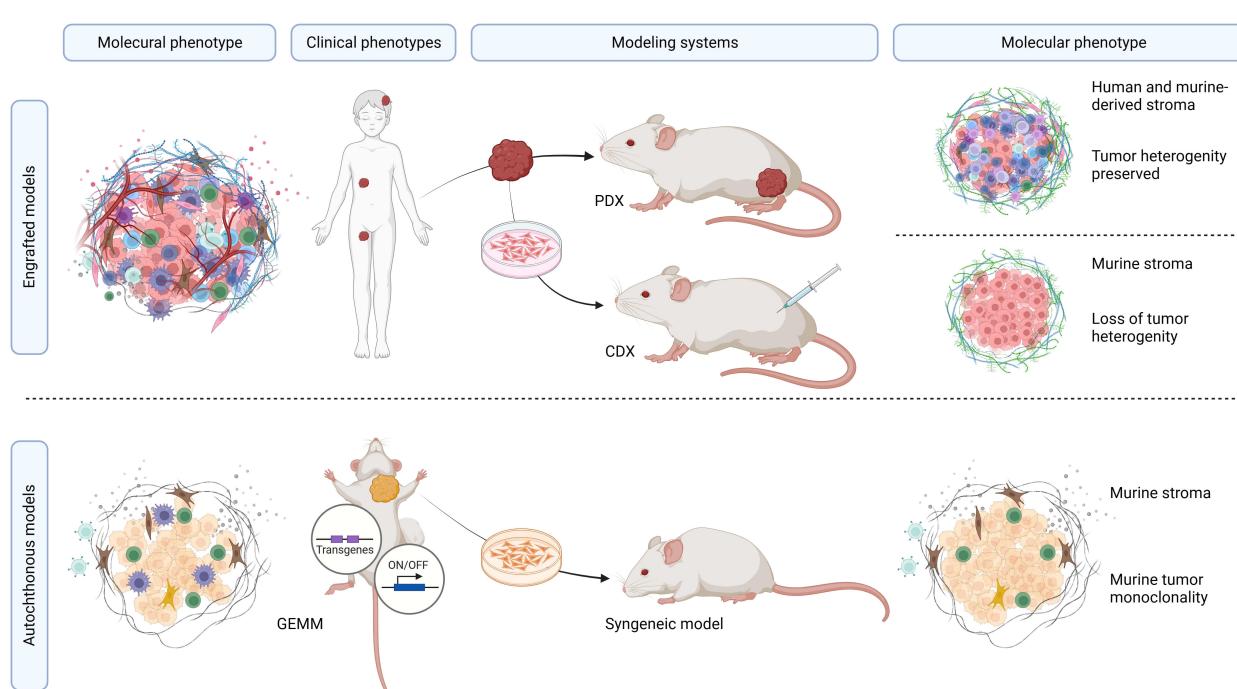


FIGURE 3

Overview of mouse model systems utilized in RMS research. Summary of preclinical RMS mouse models. The molecular tumor phenotype, in the context of the clinicopathological RMS characteristics, determines the unique features of these tumors, which can be recapitulated in both PDX and CDX models. PDX models offer the advantage of precisely mimicking the molecular, genetic, and histopathological features of the tumors while maintaining inter-patient and intra-tumor heterogeneity. CDX models, although relatively straightforward to establish and monitor, are hindered by the limited genetic diversity of human RMS cell lines within murine peritumoral microenvironments. GEMM and syngeneic models represent pure murine tumor systems with murine tumors that do not occur in humans.

