Supplementary Material
Logical Implications for Visual Question Answering Consistency

Sergio Tascon-Morales Pablo Márquez-Neila Raphael Sznitman
University of Bern
{sergio.tasconmorales, pablo.marquez, raphael.sznitman}@unibe.ch

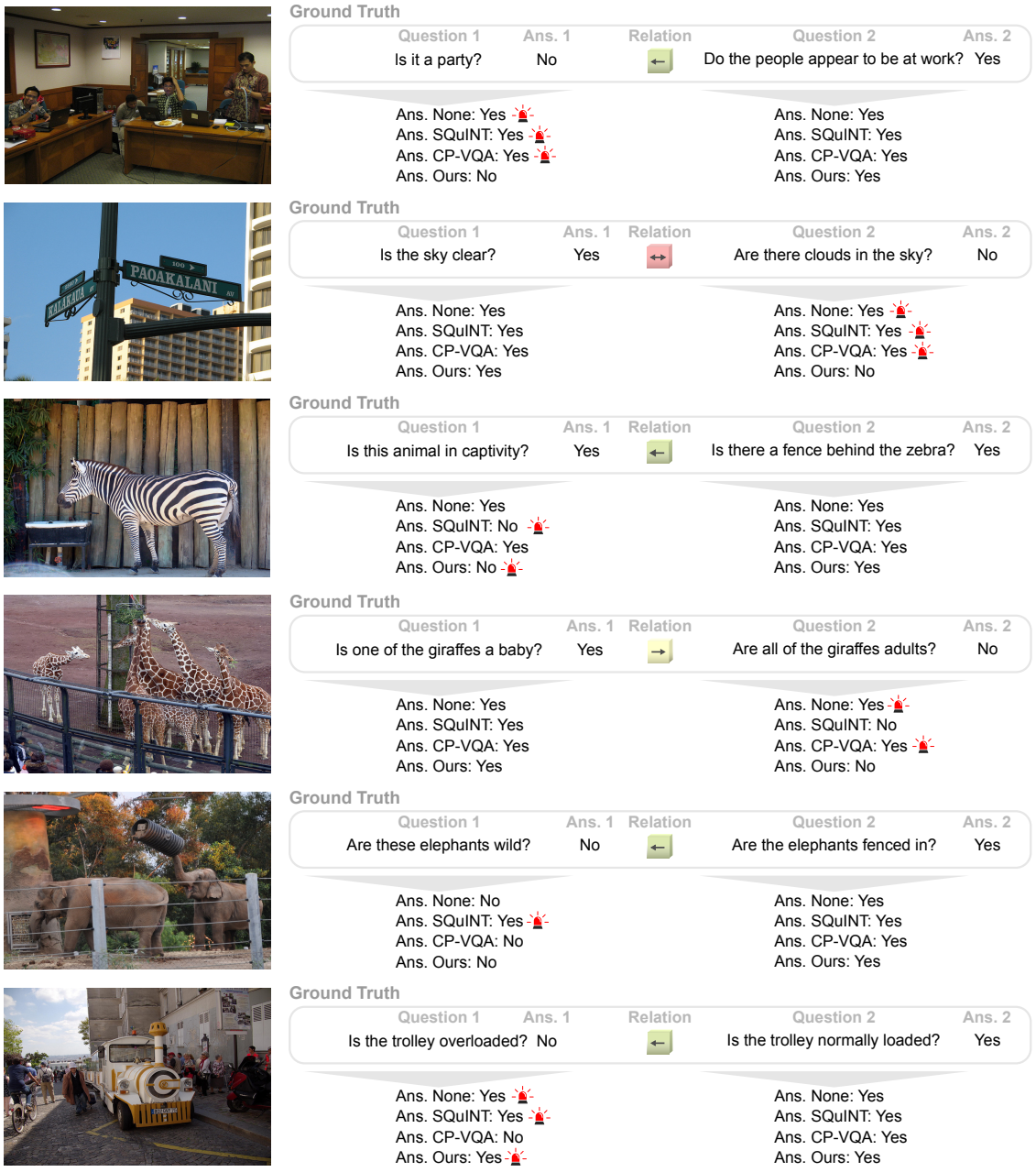

Figure 1. Additional qualitative examples from the Introspect dataset using BAN as backbone. Red siren symbols indicate inconsistent cases.

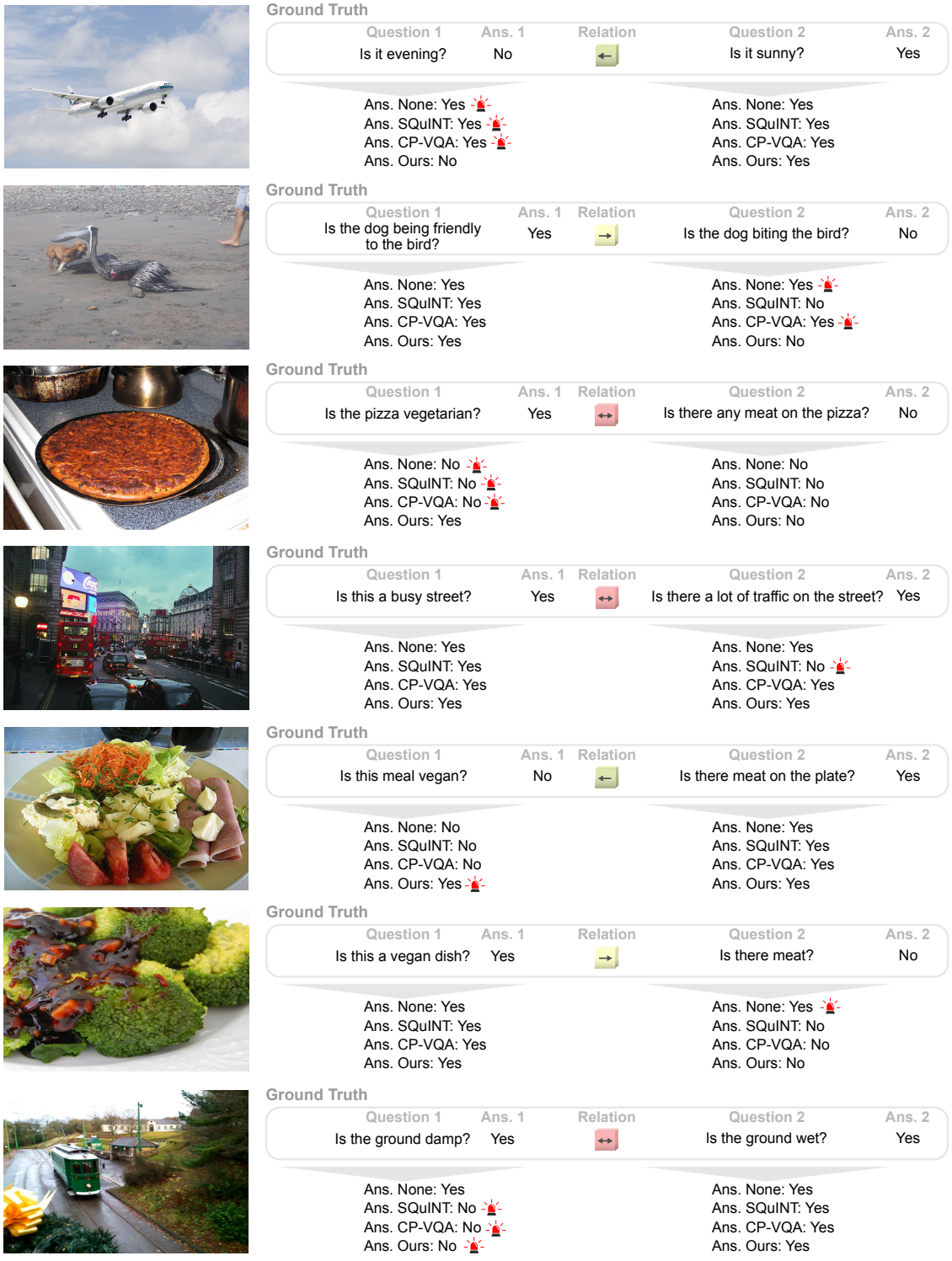

Figure 2. Additional qualitative examples from the Introspect dataset using BAN as backbone. Red siren symbols indicate inconsistent cases.

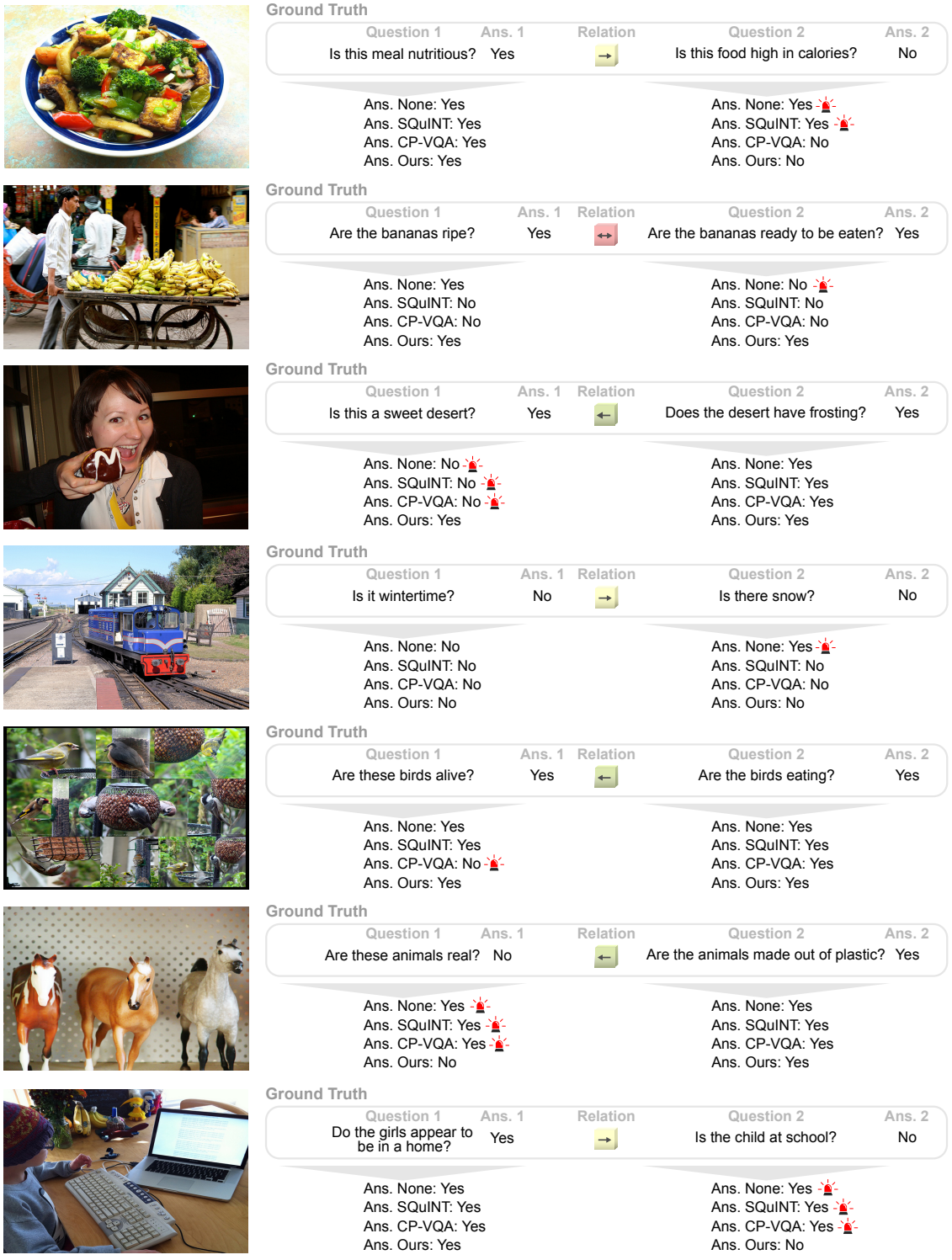

Figure 3. Additional qualitative examples from the Introspect dataset using BAN as backbone. Red siren symbols indicate inconsistent cases.

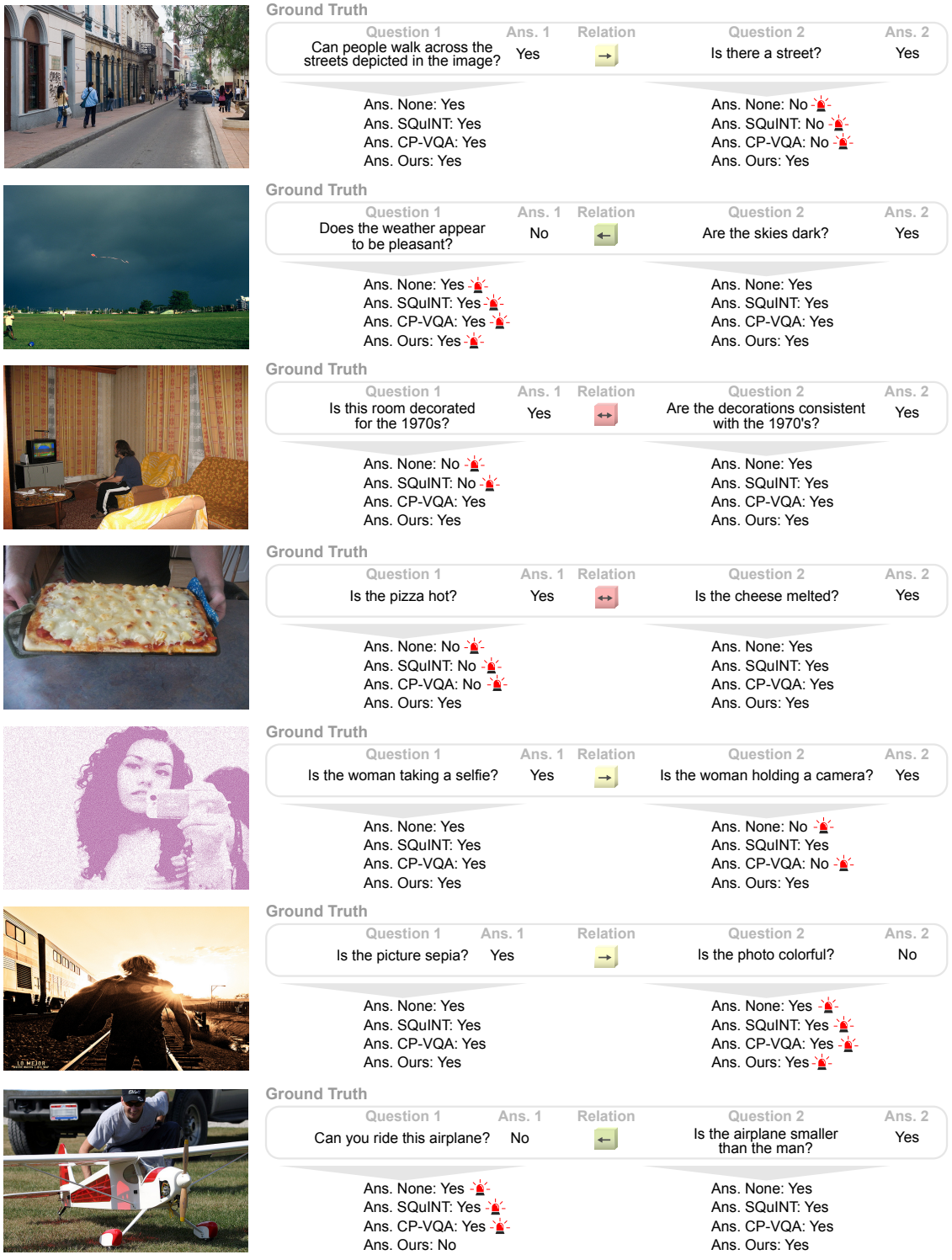

Figure 4. Additional qualitative examples from the Introspect dataset using BAN as backbone. Red siren symbols indicate inconsistent cases.

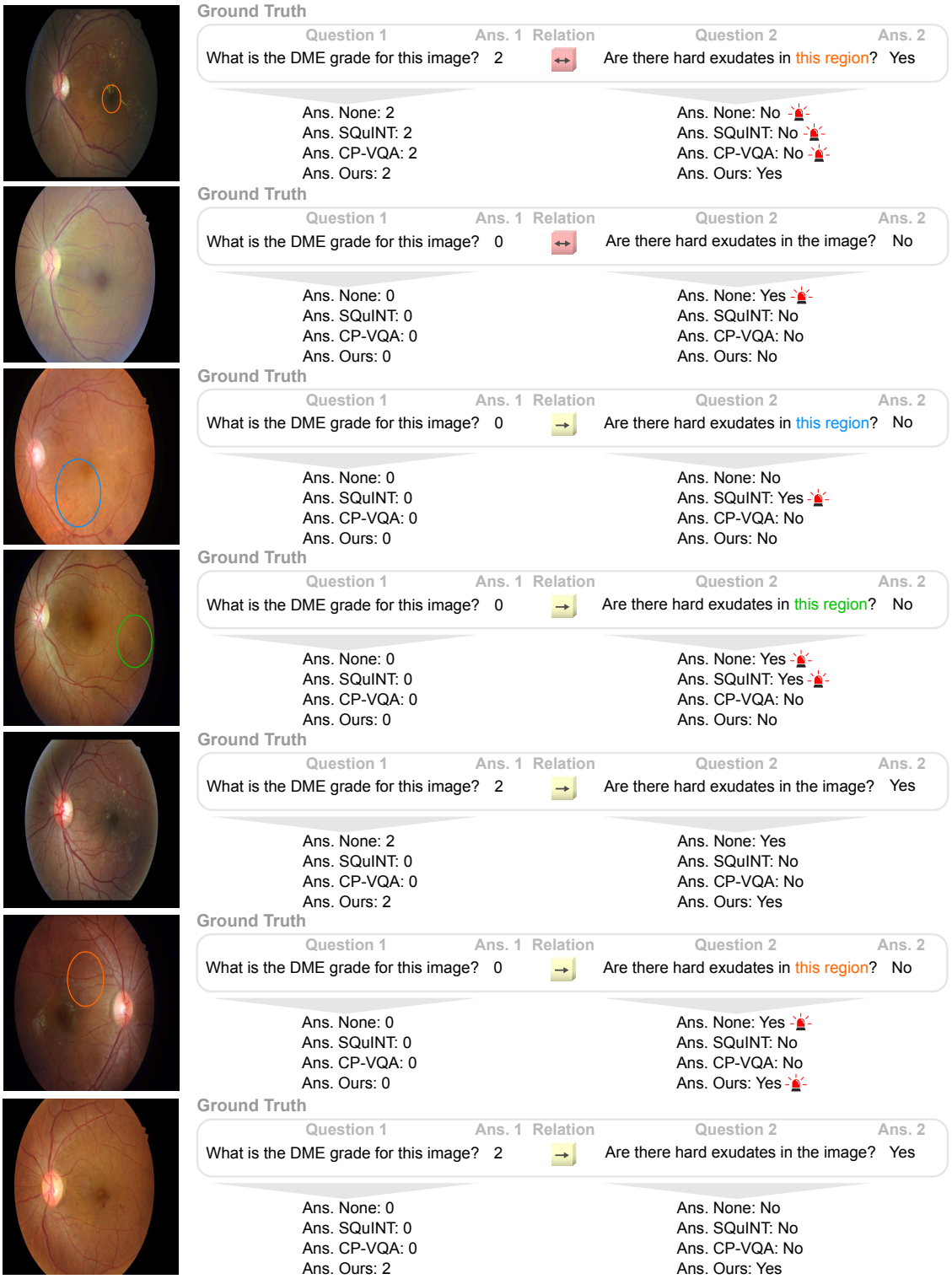

Figure 5. Additional qualitative examples from the DME dataset using MVQA as backbone. Red siren symbols indicate inconsistent cases. DME is a disease that is staged into grades (0, 1 or 2), which depend on the number of visual pathological features of the retina.

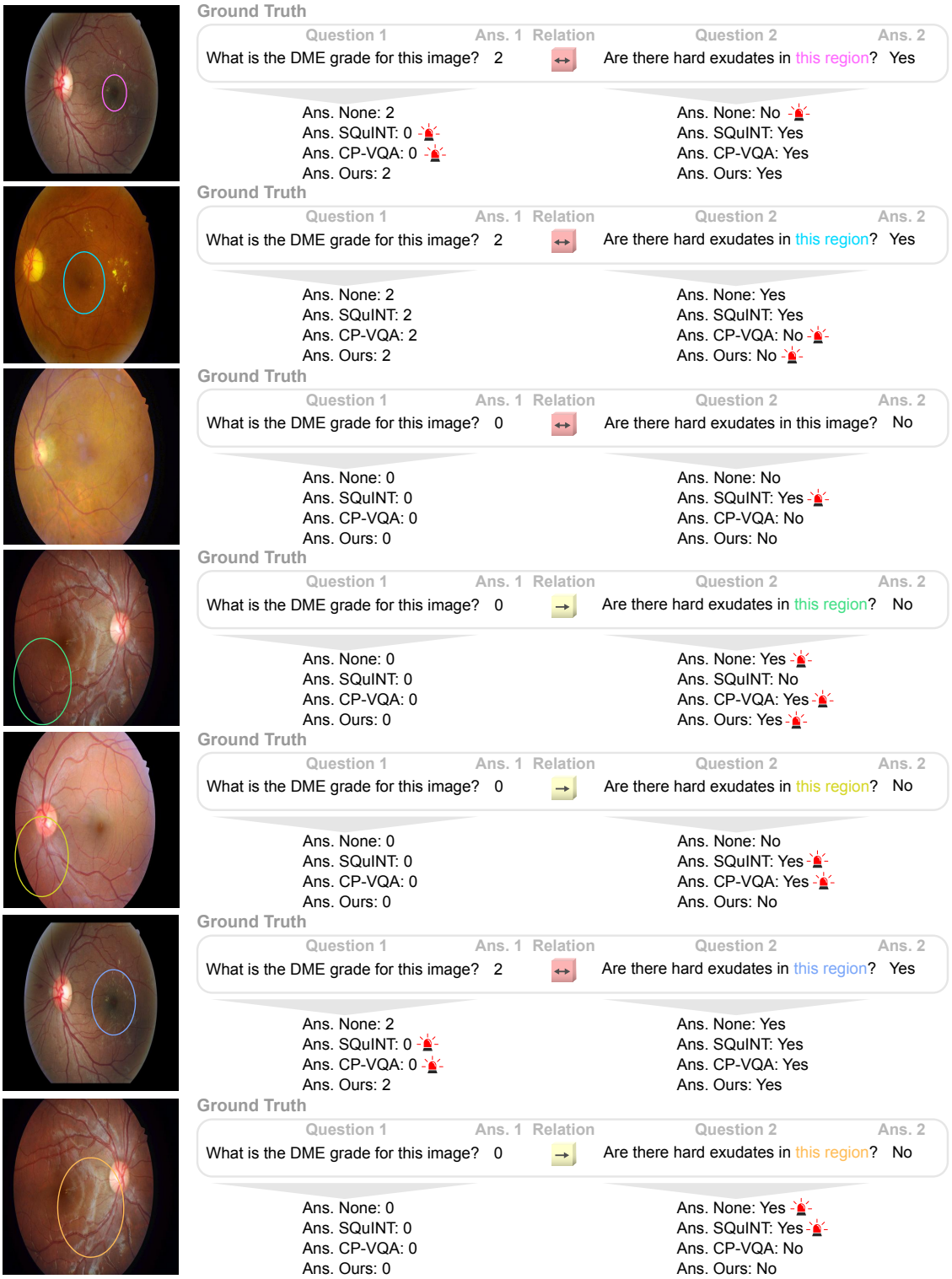

Figure 6. Additional qualitative examples from the DME dataset using MVQA as backbone. Red siren symbols indicate inconsistent cases. DME is a disease that is staged into grades (0, 1 or 2), which depend on the number of visual pathological features of the retina.