As the search for a drug begins, a natural question is “What kind of molecules can become drugs?” In 1996 Wayne Guida et al. published a review with an estimate for the number of potential molecules that could be considered as oral drug candidates. This number of molecules defines something called a molecular space. A molecular space is any set of molecules that fits defined criteria. What were our Guida’s criteria for this molecular space?

Guida defined the molecular space to include oral drugs. In general, drugs of any type need sufficient functionality to achieve a target binding energy of at least $-11 \text{ kcal/mol}$. A binding energy of that magnitude corresponds to an equilibrium dissociation constant ($K_D$ or $K_i$) of 10 nM or lower. This level of drug potency minimizes the mass of the dose that a patient receives and can also minimize side effects. A specific concern for most oral drugs is that they must be small enough to passively diffuse across membranes in the digestive system and demonstrate reasonable oral bioavailability. Lipinski put this size limit at a MW of 500. Guida put the size limit at a molecule containing no more than 30 nonhydrogen atoms, specifically C, N, O, F, P, S, Cl, Br, and I.

[Side discussion: The above paragraph describes a potential conflict. An oral drug must be large enough to generate strong binding, but the drug cannot be too big for absorption. This conflict – big enough but not too big – is a concern through most drug discovery projects.]

Guida estimated the number of possible molecules as oral drugs to be $10^{63}$. This number defines a molecular space, shown as an Euler diagram. This space is undoubtedly filled with mostly very poorly active molecules. Some regions contain compounds with low activity – hit spaces (light gray). Within some hit spaces are compounds with moderate activity – lead spaces (dark gray). Finally, some lead spaces include regions of molecules that are not only highly active but also display excellent pharmacokinetic properties and low toxicity – drug spaces (black). Note that the hit, lead, and drug spaces are drawn larger than they actually are relative to molecular space.

Potential drug space is vast. Finding a drug within this space is a needle-in-a-haystack problem. There may be some comfort, however, in the fact that for any target exist perhaps millions or billions of molecules that could be approved as drugs. The idea that a drug group must find the one best molecule is almost certainly false. The job of discovery team is to find a molecule that is good enough, not necessary the best.
Original reference: